

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Puliluker Type of Search 1 Vendors and cost where applicable
Searcher Phone #: 308-4740 NA Sequence (#) 2 STN _____
Searcher Location: 1507 AA Sequence (#) 2 Dialog _____
Date Searcher Picked Up: 6-29 Structure (#) _____ Questel/Orbit _____
Date Completed: 7-2-01 Bibliographic _____ Dr. Link _____
Searcher Prep & Review Time: 5 Litigation _____ Lexis/Nexis 01, 02
Clerical Prep Time: _____ Fulltext _____ Sequence Systems _____
Online Time: 5 Patent Family _____ WWW/Internet _____
Other _____ Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2001, 23:49:58 ; Search time 1372.68 Seconds
(without alignments)
12147.648 Million cell updates/sec

Title: US-09-457-066-1
Perfect score: 1764
Sequence: 1 attatgtggaactaccctg.....cttgtgtcgtgctgatagga 1764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_est4:*
 - 5: gb_est5:*
 - 6: gb_est6:*
 - 7: gb_est7:*
 - 8: gb_est8:*
 - 9: gb_est9:*
 - 10: gb_est10:*
 - 11: gb_est11:*
 - 12: gb_est12:*
 - 13: gb_est13:*
 - 14: gb_est14:*
 - 15: gb_est15:*
 - 16: gb_est16:*
 - 17: gb_est17:*
 - 18: gb_est18:*
 - 19: gb_est19:*
 - 20: gb_est20:*
 - 21: gb_est21:*
 - 22: gb_est22:*
 - 23: gb_est23:*
 - 24: gb_est24:*
 - 25: gb_est25:*
 - 26: gb_est26:*
 - 27: gb_est27:*
 - 28: gb_est28:*
 - 29: gb_est29:*
 - 30: gb_est30:*
 - 31: gb_est31:*
 - 32: gb_est32:*
 - 33: gb_est33:*
 - 34: gb_est34:*
 - 35: gb_est35:*
 - 36: gb_est36:*
 - 37: gb_est37:*
 - 38: gb_est38:*
 - 39: gb_est39:*
 - 40: gb_est40:*
 - 41: gb_est41:*
 - 42: gb_est42:*
 - 43: gb_est43:*
 - 44: gb_est44:*
 - 45: gb_est45:*
 - 46: gb_est46:*
 - 47: gb_est47:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	28.3	696	153	BG430400
2	486.4	27.6	500	166	BE326924
3	482.2	27.3	523	142	BE958470
4	478.8	27.1	851	144	BF102859
5	465.8	26.4	565	141	BE884591
6	464.6	26.3	950	143	BF031624
7	459.4	26.0	874	168	BF670092
8	455.6	25.8	820	168	BF697294
9	444	25.2	556	155	BG609411
10	441.4	25.0	655	153	BG432570
11	432.2	24.5	877	145	BF137533
12	411	23.3	412	190	W21436
13	407.8	23.1	564	145	BF151355
14	400	22.7	983	174	BG112388
15	396.2	22.5	889	168	BF698737
16	394.4	22.4	425	11	AA759138
17	390	22.1	467	105	AL040028
18	384.4	21.8	523	143	BF021679
19	380	21.5	968	174	BG118707
20	378.8	21.5	511	113	AW210331
21	371	21.0	764	150	BF541705
22	355.6	20.2	886	146	BF217515
23	350	19.8	1003	174	BG113879
24	349.2	19.8	715	141	BE879725
25	336	19.0	631	122	AW958887
26	334.6	19.0	393	22	A1567614
27	333	18.9	609	144	BF133918
28	330.8	18.8	447	143	BF011835
29	324	18.4	910	175	BG243001
30	312	17.7	400	105	AL364180
31	311.6	17.7	408	166	BE304156
32	303.4	17.2	354	20	A1470494
33	302.4	17.1	304	105	AL047637
34	287	16.3	1029	174	BG173847
35	286	16.2	423	105	AL042106
36	279.4	15.8	865	143	BF032143
37	256.4	14.5	864	174	BF137524
38	254.8	14.4	346	174	BG147080
39	251	14.2	834	144	BF130765
40	243	13.8	243	111	AW052200
41	226.4	12.8	324	15	A1020581
42	226.4	12.8	892	146	BF247297
43	226	12.8	226	6	AA358262
44	225.8	12.8	832	144	BF130122
45	216.8	12.3	279	147	BF327185

ALIGNMENTS

RESULT	1
LOCUS	BG430400
DEFINITION	602502113F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4615755 5', mRNA EST 14-MAR-2001
ACCESSION	BG430400
VERSION	BG430400.1 GI:13336906
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 696)
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1995) Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCMI369 row: i column: 04
High quality sequence stop: 559.
Location/Qualifiers
source
1. 696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4615755"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (fl phage-resistant)"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccattatgcc); Site_2: SfiI (ggccattatgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGGCCATATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

FEATURES

BASE COUNT	201 a	128 c	160 g	207 t
ORIGIN				
Query Match	28.38;	Score	500;	DB 153; Length 696;
Best Local Similarity	93.78;	Pred. No.	9.5e-126;	
Matches	579;	Conservative	0;	Mismatches 25; Indels 14; Gaps 5;
QY	1153	gagtgactgtgtgcagaggagcacagagagatagcgcacacacacagcgtc	1212	
Db	1	GAGTGTGACGTGTGTGTCGAGAGGAGGACACAGAGATAGCCGCATCACACACAGAGGTC	60	
QY	1213	ttgcccagagctgtgcaggtlygcagtggtgatcttattagagaacgtatgcgtatc	1272	
Db	61	TTGCCCAGAGCTGTGCAGT----GCAGTGTGCTGATTCATTAGAGAACGTATGCGTTATC	116	
QY	1273	tcacatcctaactcagttgttgccttcacagacatttcattcattcagatttacagtcca	1332	
Db	117	TCCATCCTTAATCTCAGTTGTTGCTTCAAGGACGCTTTCATCTTCAGGATTTACAGTGCA	176	
QY	1333	tctgaaagagagacatcaaacagaattagagttgtgcaacagctcttttgaagagag	1392	
Db	177	TTCTGAAGAGGAGACATCAACAGAAATAGGAGTTGTGCAACAGCTCTTTGAGAGGAG	236	
QY	1393	gcctaaagagagagaaaaaggtctcaatcgtagg-aaagaaaaataatgtgtatcaa	1451	
Db	237	GCCTAAGGACAGGAGAAAGGTCTTCAATCGTGGCAAGAAATTAATTTGTTATTAA	296	
QY	1452	atagatcaccagctagtttcagagttaccatgtacctattccactagctgggttctgtat	1511	
Db	297	ATAGATCACCAGCTAGTTTTCAGAGTTACCATGTACCATGTCCACTAGCTGGTTCGTAT	356	
QY	1512	ttcagttcttcgatacagcttagggtaatttcagtcagacagaaaaaaactgtgaagtga	1571	
Db	357	TTCAGTCTTTTCGATACGGCTTAGGGTAATGTCAGTACAGAAAAAACTGTGCAAGTGA	416	
QY	1572	gcacctgattccgtgcttgcttaactcaagctcctcctggcctcaaatcgta	1631	
Db	417	GCACCTGATTCCGTTGCTTGAACCTTAACCTCAATGCTCCTGGCCCTAAATTCGTA	476	
QY	1632	taaaatcgtgatt	1688	
Db	477	TAAATCTGGA----TTTTTCTTTTGGGCTCATATTTCACATTTGTAAAGCCAGAACAT	532	
QY	1689	attctatgtactacaacacctggttttttaaa--aaggaaactatgttgcataaataaact	1746	

Qy	1438	aaatgtgtattaaatagatcaccagctagttcagaagttaccatgtacgtattaccacta	1497
Db	320	AAATGTTGTATTAATATAGATCACCAGCTAGTTTTCAGAGTTTACCATGTACCTATTCACCTA	261
Qy	1498	gctgggtctctatttcagttctcttcgatacgcgttaggttaatgtcagtacagaaaaa	1557
Db	260	GCTGGGTTCTGTATTTTCAGTCTCTTTCATACGGCTTAGGGTAATGTTCAGTACAGGAAAAA	201
Qy	1558	aactgtgcaagtgcacctgattccgttgcttcctgttaactctaaagctccatgtccgtg	1617
Db	200	AACTGTGCAAGTGAGCACCTGATTTCGTTGCCTTGCTTAAGTCTAAAGCTCCATGTCCCTG	141
Qy	1618	ggcctaaatcgtatataaaatcgtgatttttttttttttttttttttttttttttttttttttt	1677
Db	140	GGCCTAATAATCGTATATAATCTGGA--TTTTTTTTTTTTTTTTTTTTTCTCATATTCATATG	83
Qy	1678	taaaccaagaaccttatctatgactacaaacctgctttttaaaaggaaactatgttgctatg	1737
Db	82	TAAACCAAGAACATCTATGTACTACAAACCTGGTTTTTAAAGGAACATGTGTGCTATG	23
Qy	1738	aataaaactgtgtcgtgctga	1759
Db	22	AATTAACACTGTGTCATGCTGA	1
RESULT	3		
BE958470			
LOCUS	601644787F1	N1H_MGC_56	Homo sapiens cDNA clone IMAGE:3930045 5',
DEFINITION	BE958470	mRNA	EST 04-OCT-2000
ACCESSION	BE958470	mRNA sequence.	
VERSION	BE958470.1	GI:10569175	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 523)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: CLONETECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: LLCM763 row: m column: 22		
	High quality sequence start: 2		
	High quality sequence stop: 513.		
FEATURES	Location/Qualifiers		
source	1..523		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3930045"		
	/clone_lib="NIH_MGC_56"		
	/tissue_type="primitive neuroectoderm"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site:1		
	SfiI (ggccgctcgcc); Site:2: SfiI (ggccattatggcc);		
	Double-stranded cDNA was prepared from cell line RNA. 5'		
	and 3' adaptors were used in cloning as follows: 5'		
	adaptor sequence: 5'-CAGCGCCATTATGCC-3' and 3' adaptor		
	sequence: 5'-ATTCTAGAGCGAGCGCCGACATG-dT(30)BN-3'		
	(where B = A, C, or G and N = A, C, G, or T). Average		
	insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies		
	contained inserts by PCR. This library was enriched for		
	full-length clones and was constructed by Clontech		
	Laboratories (Palo Alto, CA)."		
BASE COUNT	147 a 118 c 129 g 128 t		1 others

ORIGIN

Query Match 27.3%; Score 482.2; DB 142; Length 523;
Best Local Similarity 98.7%; Pred. No. 6.6e-121;
Matches 517; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
QY 677 tgcacccttcagcttgccactggaactgcttaataatgctataactgctttagta 736
Db 1 TGTACCCCTTCAGC-TTGCCACTGGACCTGCTTAATAATGCTATAACTGCTTTAGTA 59
QY 737 ccttggaagacattatcagatattctggaagagagatggcagttggacttagaagtc 796
Db 60 CCTTGAAGACCTTATTTCGATATCTTGAACCCAGAGAGATGGCAG-TGGACTTGAAGATC 118
QY 797 tatatagccaacttggaactcttggaagctttgtttttggaagaaataccagag 856
Db 119 TATATAGGCCAACTTGGCAACTTCTTGGCAAGCTTCTGTTCGGAAGAAATCCAGAG 178
QY 857 tgggtgatctgaacctcttaacagagaggttaagattatcacagctcacacctcgtaact 916
Db 179 TGGTGGATCTGAACCTTCTTAACAGAGAGGTAGATATATACAGCTGCACACCTTCGTAAC 238
QY 917 tctcagtgctccataaggggaagaaactaaagaaagaccgataccattttctgccaagttgtc 976
Db 239 TCTCAGTGTCCATAAGGGAAGAACTAAACAGAACCCGATACCATTTCTGCCAGGTGTC 298
QY 977 tctcagttaaacgctgtggtggaactgtgcctgtgtgtctccacaattgcaatgaatgtc 1036
Db 299 TCTCGTTAAACGCTGTGTGGGAAGTGTGCTGTGTCTCCCAATTCGAATGAATGTC 358
QY 1037 aatgtgtcccaagcaaaagttaactaaataacacagagctcttcagttgagaccaga 1096
Db 359 AATGTGTCCCAAGCAAGTACTAAATAATACACAGAGTCTTCAGTTGAGACCAAGAA 418
QY 1097 ccggtgtcaggggattgcacaactcaactcaccgagctggtggtggtggtggtggtggt 1156
Db 419 CCGGTGTACAGGATTTGCACAACTCACTCACCAGCTGG-CCTGGAGACCATGAGAGT 477
QY 1157 gtgactgtgtgagagagggacagagaggtagtagccgcatcac 1200
Db 478 GTGACTGTGTGTCAGAGGGAGCAGAGGAGTAGCCGAATCAC 521

RESULT 4

BF102859 851 bp mRNA EST 19-OCT-2000
LOCUS 601646827F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4073095 5',
DEFINITION mRNA sequence.
ACCESSION BF102859
VERSION BF102859.1 GI:10885385
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 851)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10CM921 row: f column: 08
High quality sequence stop: 529.
Location/Qualifiers
1. .851

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4073095"
/clone_lib="NIH_MGC_60"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggccattatggc)
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dT(30)BN-3',
(where B = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 234 a 169 c 238 g 210 t
ORIGIN
Query Match 27.1%; Score 478.8; DB 144; Length 851;
Best Local Similarity 96.6%; Pred. No. 6.3e-120;
Matches 543; Conservative 0; Mismatches 12; Indels 7; Gaps 5;
QY 639 tgcatacccaattccacagaagctgtgagctctcagtgctaccccttcagcttgc 698
Db 8 TGTATGCCACAATTCAGAAAGCTGTGAGTCTTCACTGCTACCCCTTTCAGC-TTGC 66
QY 699 actggacctcttaataatgctataactgcttttagtaccttggaagaccttattcgata 758
Db 67 ACTGGACCTGCTTAATAATGCTTAACCTGCTTTAGTACCTTGGAGAGCTTATTTCGATA 126
QY 759 tcttgaaaccagagagatggcagttggacttagaagatctatagagcaacttggcaact 818
Db 127 TCTTGAACCCAGAGAGATGGCAGTTGGACTTAGAAGATCTATATAGSCCAACTTGGCAACT 186
QY 819 tcttgcaagagctttgtttttggaagaaatccagagtggtggtctgaaaccttctaac 878
Db 187 TCTTGGCAAGGCTTT--GTTTTGGAAGAAATCCAGAGTGGTGGATCTGAACCTTCTAAC 244
QY 879 agagaggttaagattatcacagctcacacctcgtaactctcagtgctccataagaga 938
Db 245 AGAGAGGTAAAGATTATACAGCTGCACACCTCGTAACCTCTCAGTGTCCATTAAGGAAGA 304
QY 939 actaaagagaaccgataccattttctggccaggtgtgtctcctgggttaaacgctgtgtg 998
Db 305 ACTAAAGAGAACCGATACCATTTCTGCGCAGGTTGTCTCTGTTTAAACGCTGTGGTGG 364
QY 999 gaactgtcctgtgtgtctccacaattgcaatgaatgtcaatgtgtcccaagcaagttac 1058
Db 365 GAACGTGTGCTGTG-TGCTCTCCACAATTCGAATGAATGTCAATGTCTCCCAAGCAAGTTAC 423
QY 1059 taaaaataccacagagctcttcagttgagaccaaagaccggtgtcagggattacaa 1118
Db 424 TAAAAATATACCACGAGGTCTTCAGTTGAGACCAAGACCCTGTCTCAGGGGA-TGCACAA 482
QY 1119 atcaactcaccagctgtgcccctggagcaccatgagagtggtgactgtgtgagagggag 1178
Db 483 ATCACTCACCAGCTGGGCC--TGAGCACCATGAGAGGTGTGACTGTGTGTGCGAGGAGC 540
QY 1179 cacaggaggatagccgcacac 1200
Db 541 CACAGGGGATAGCCGATTAAC 562
RESULT 5
BE884591
LOCUS BE884591 565 bp mRNA EST 20-OCT-2000
DEFINITION 601506290F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3907918 5',
mRNA sequence.
ACCESSION BE884591

BE884591.1 GI:10333367
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 565)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM9719 row: c column: 23
 High quality sequence stop: 565.
 Location/Qualifiers
 1..565
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3907918"
 /clone_lib="NIH_MGC_71"
 /tissue_type="lelomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Not I;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
 Average insert size 2.1 kb."
 BASE COUNT 182 a 101 c 125 g 157 t
 ORIGIN

 Query Match 26.4%; Score 465.8; DB 141; Length 565;
 Best Local Similarity 99.2%; Pred. No. 2.le-116;
 Matches 489; Conservative 0; Mismatches 2; Gaps 2

 QY 268 aacggagtcacagatcctcagcatgagagaattactgtctactaatggaagtatt 327
 |||
 Db 73 AAGGGAGTACAGATCCTCAGCATGAGAAATATTACTGTGTCACACATGGAAGTATT 132
 |||
 QY 328 cacagcccaagggttctcctcactactatccaaagaaatacggctctggtatggagattagta 387
 |||
 Db 133 CACAGCCCAAGGTTTCTCTACTATTATCCAGAAATACGGTCTTGGTATGGAGATTAGTA 192
 |||
 QY 388 gcagtagagagaaaatgtagtatacacttaacgtttgatgaaagattgggcttgaagac 447
 |||
 Db 193 GCAGTAGAGGAAATGATGCGATACAACTTACGTTTGTATGATGAAAGATTTGGGCTTGAGAC 252
 |||
 QY 448 ccagaagatgacatgacgaagtattgtttagaagttgaggaacccagtgatggaact 507
 |||
 Db 253 CCAGAAGATGACATATGCAAGTATGATTTGTAGAAGTTGAGGAACCCAGTGATGGAAC 312
 |||
 QY 508 atattaggcgctgggtgtgtctgtgactgtaccagagaaacagatttctaagggaat 567
 |||
 Db 313 ATATTAGGCGCTGGTGTGGTCTGGTACTGTACAGAGAAACAGATTTCTTAAGGAAAT 372
 |||
 QY 568 caaattaggataaagatttgtagtataatttctcctctgaaccagggttctgcac 627
 |||
 Db 373 CAAATTAGGATAAGATTTGATCTGATGAATATTTTCCTTCTGAACACAGGGTTCTGCATC 432
 |||
 QY 628 cactacaacatgtcatgccacaattccagaagctgtgagtccttcagtactaccccc 687
 |||
 Db 433 CACTACAACATGTCATGCGCAATTCACAGAAGCTGTGAGTCCTTCAGTCTACCCCT 492
 |||
 QY 688 tcagctttgccactggaactgcttaataatgcat-aactgccttttagtaccttg-gaag 745
 |||
 Db 493 TCAGCTTTGCCACTGGACCTGCTTAAATATGCTAATAACTGGCTTTAGTACTTCGGAAG 552
 |||
 QY 746 accttattcgata 758

```

Db      553  ACCTATTTCGATA 565

RESULT 6
BF031624
LOCUS   601558104F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3827760 5',
DEFINITION
mRNA sequence.
ACCESSION BF031624
VERSION   BF031624.1 GI:10739336
KEYWORDS EST.
SOURCE    human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM497 row: h column: 01
          High quality sequence stop: 415.

FEATURES
        source
            Location/Qualifiers
                1..950
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:3827760"
                    /clone_lib="NIH_MGC_58"
                    /tissue_type="hypernephroma"
                    /lab_host="DH10B (T1 phage-resistant)"
                    /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site1:
                    SfiI (ggcgcttcggcc); Site2: SfiI (ggcgcttcggcc);
                    Double-stranded cDNA was prepared from cell line RNA. 5'
                    and 3' adaptors were used in cloning as follows: 5'
                    adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
                    sequence: 5'-ATTCTAGAGCGAGCGCGCATG-3' (30)BN-3'
                    (where B = A, C, or G and N = A, C, G, or T). Average
                    insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies
                    contained inserts by PCR. This library was enriched for
                    full-length clones and was constructed by Clontech
                    Laboratories (Palo Alto, CA)."
                288 a 309 c 211 g 142 t

BASE COUNT
ORIGIN

```

		Query Match	26.3%;	Score 464.6;	DB 143;	Length 950;
		Best Local Similarity	85.6%;	Pred. No. 5e-116;		
		Matches 578;	Conservative	0;	Mismatches 84;	Indels 13; Gaps 5;
Qy	902	gcacacctcgtaactcttcagttgccataagggaagaactaaagagaccatgataccattt	961			
Dd	1	GCACACCTCGTAATTCTCACTGTCCATAGGGGAGAAGAACTAAAGAGAACCCGATACCATT	60			
Qy	962	tctgtggcagggtgtctctcgtgttaaaccgctgtgtgggaactgtgcctgttgtctccaca	1021			
Dd	61	TCTGCCCAGGTTGTCTCTCGTGTAAACGCTGTGGTGGAAGTGTGCCTGTGCTCCACA	120			
Qy	1022	attgcaatgaatgtcaatgtgtcccagcaagttactaaaaataaccacagagtccttc	1081			
Dd	121	ATTCCAATGAATGTCAATGTGTCCTCCAGCAAGTTACTAAAAAATATCACGAGGTCCTTC	180			
Qy	1082	agtgagaccaaaaagaccoggtgtcgaggggatgtgcacaatcatcaccgacgtggccctgg	1141			
Dd	181	AGTTGAGACCAAAGACCGGTGTGAGGGGATGTGCACAAATCATCTCCGAGCATGGCCCTGG	240			


```
ACCESSION   BF697294
VERSION     BF697294.1  GI:11982702
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 820)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: CLONETECH Laboratories, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC alone distribution information can be
            found through the I.M.A.G.E. Consortium/ILLNL at:
            http://image.llnl.gov
            Plate: L1CM1123 row: o column: 08
            High quality sequence stop: 462.

FEATURES             source
     1..820
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:4286815"
     /clone_lib="NTH MGC_56"
     /tissue_type="primitive neuroectoderm"
     /lab_host="DH10B (T1 phage-resistant)"
     /note="Organ: brain; Vector: pONR-LIB (Clontech); Site: 1:
     SfiI (ggccgcccggcc); Site: 2: SfiI (ggccattatggcc);
     Double-stranded cDNA was prepared from cell line RNA.
     and 3' adaptors were used in cloning as follows: 5'
     adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
     sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-df(30)BN-3'
     (where B = A, C, or G and N = A, C, G, or T). Average
     insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
     contained inserts by PCR. This library was enriched for
     full-length clones and was constructed by Clontech
     Laboratories (Palo Alto, CA)."
     BASE COUNT      221 a 179 c 195 g 225 t
     ORIGIN

     Query Match      25.8%; Score 455.6; DB 168; Length 820;
     Best Local Similarity 93.6%; Pred. No. 1.4e-113;
     Matches 498; Conservative 1; Mismatches 25; Indels 8; Gaps 2;

QY 1235 ygcagtggctgattattagagaacgtatgcgttatctccatccttaattcagttgtt 1294
      :|||||
Db 13  TGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGT 72

QY 1295 tgcctcaaggaccttcattcattcagagattacagtgcatcttgaaggagagacatcaaa 1354
      :|||||
Db 73  TGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATCTCTGAAGAGGAGACATCAAA 132

QY 1355 cagaattaggagttgtgaacagctcttttgagaggagccttaaggagagagagaaagg 1414
      :|||||
Db 133 CAGAAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCGCTTAAGGAGGAGAAAGG 192

QY 1415 tottcaatcgtggaagaaaaataaattgtgtattataatagatcacccagctagtttcaga 1474
      :|||||
Db 193 TCTTCAATCGTGGAAAGAAAATTAATGTGTATTAAATAGATCACCAGCTAGTTTCAGA 252

QY 1475 gttaccatgacgtattccactagctggttctgtatttcagcttcttgcagcgtta 1534
      :|||||
Db 253 GTTACCATGTACGATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTTCGATACGCGTTA 312

QY 1535 gggttaattcagtcagagaaaaaactgtgaagtgcacacacctgattccgttgccttgc 1594
      :|||||
Db 313 GGGTAATGTACGACAGGAAAAAACCCTGTGCAAGTGAGCACCTGATTCGTTGCGCTTGCT 372

QY 1595 taactctaaagctccatgtcctgggctaaacgtataaaatcgtataaaatcgtggtttttt 1654
      :|||||
Db 373 TAACTCTAAAGCTCCATGTCTCTGGCGCTAAATCGTATAAAATCTGGA-----TTTFTT 426

QY 1655 tttttgctcatattcacatattgaaacagaaacattctatgtactacaaacacctgggttt 1714
      :|||||
Db 427 CTTTGGGCCCATTTCCCTTTTGTAAACCAGAACTTTCTATGTCTTAACAACACCTGGTTT 486

QY 1715 taaaaggaa--ctatgttctgatgaattaaacctgtgtgtgtgtgtatagga 1764
      :|||||
Db 487 TCAACACGGGACCCCTATGTGGCTATGAATTAACCTGGTGTGCTGTGTATAGGA 538

RESULT 9
BG609411 556 bp mRNA EST 17-APR-2001
LOCUS 323251 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG609411
ACCESSION BG609411
VERSION 1 GI:13659390
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 556)
AUTHORS Fahrrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keeler,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt.trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR Primers
            FORWARD: AGGAACACGCTATGACCAT
            BACKWARD: GTTTCACGTCACGACG
            Plate: 98 row: P column: 8
            Seq primer: ATTAGGTGACACTAG.
FEATURES             Location/Qualifiers
     1..556
     /organism="Sus scrofa"
     /db_xref="taxon:9823"
     /clone_lib="MARC 1PIG"
     /tissue_type="pooled"
     /lab_host="DH10B"
     /note="Vector: pCMV SPORT6; Site: 1: XbaI; Site: 2: XhoI;
            Library made from pooled tissue from day 11, 13, 15, 20,
            and 30 embryos."
     BASE COUNT      139 a 153 c 141 g 123 t
     ORIGIN

     Query Match      25.2%; Score 444; DB 155; Length 556;
     Best Local Similarity 87.4%; Pred. No. 1.9e-110;
     Matches 486; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 615 agggttctgcattccattcaaacattgtcgtccacattcacagaagctgtgagctctc 674
      :|||||
Db 1  AGGTTCTGCATCCATCAACATCGTCACACCAATTCACAGAACTGTGAGTCCCTC 60

QY 675 agtgcaccoccttcagcttggcactggacctgcttcaataatgctataactgccttag 734
      :|||||
Db 61 ACTGCTACCGCTGCAGCTCTGCCCTGGACCTGCTTAACAATGCTCTACTGCTTAG 120

QY 735 taccttggagaccttattogatatttgaaccagagagatggcagttggacttagaaga 794
      :|||||
```


Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LAM9243 row: O column: 08
High quality sequence stop: 677.

FEATURES

source
Location/Qualifiers
1..877
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4008799"
/clone_lib="NCI-CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
221 a 220 c 233 g 203 t

Query Match 24.5%; Score 432.2; DB 145; Length 877;
Best Local Similarity 79.0%; Pred. No. 3.6e-107;
Matches 552; Conservative 0; Mismatches 143; Indels 4; Gaps 3;
QY 272 gagtacaagatcagtcagagagatattactgtctactaagagatgagagatcaca 331
|||||
Db 30 GAGTGAAGATCCCGGATGAGAGAGTGTCTACTATATCTGTATGTTGGAGGATCCACA 89
QY 332 gcccaagggttctctactattcccaagaatacgcgtctgtgtgtagagattagtagcag 391
|||||
Db 90 GCCGGAAGTTTCTCTACATACCCCAAGAAATATGTGTGTGTGGAGATTAGTTGCAG 149
QY 392 tagagagaatgtatgatacaactacgtttgtatgaagatttggtgctgaagaccag 451
|||||
Db 150 TAGATGAAATGTCCGGATCCAGCTGACATTTGTATGAGAGATTGGCTGGAAGATCCAG 209
QY 452 aagatgacatattcagatgattgttagaagttgagagaccagtgatggaactatat 511
|||||
Db 210 AGAGCATCTATGCAAGTATGATTTGTAGAGTTGAGGAGCCAGTGTGGAAGTGT 269
QY 512 tagggcgtggt 571
|||||
Db 270 TAGGACGCTGT 329
QY 572 ttaggataagattgtatgtatgaatatatttctctctgaaccagggttctgcatccact 631
Db 330 TCAGGATAAGATTGTATCTGATGATGATTTTCCATCTGAACCCGGATTCTGCATCCACT 389
QY 632 acaacattgtcatgccacaattcacagaagctgtgagctctcagtgctaccccttcag 691
|||||
Db 390 ACAGTATTATCATGCCACAGTCACAGAACACAGAGTCTCTCGGTGTGCCCCCTTCAT 449
QY 692 cttgcccactggacctgcttaataatgtataactcctttagtaccttgg-aagacctt 750
Db 450 CTTTGTCTATGGACCTGCTCAACACGCTGTGACTGCTCTCAGTACCTTGGAAAGAGCTG 509
QY 751 attcgatatcttgaaccagagatggcaggttggagctttagaagatctatatagggccaact 810
Db 510 ATTCCGTACTAGACGCAGATCGATGGCAGGTGGACTTGGACAGCCCTTACAGCCCAACA 569
QY 811 tggcaactcttggcaagccttttggtaagaagaaatccagagtggtgagatctgaac 870
Db 570 TGGCAGCTTCTTGGGCAAGGTTTCCCTGTATGGGAACACACACAGCAGGAGTGAATCTCAAT 629
QY 871 ctcttaacag--aggaggttaagattatacagctgcacactcgtgaactctcagtgcca 928
Db 630 CTCCTAAAGGGAGGAGTAAACACTCTACAGTGGCACACCCCGGAATTCCTCAGGTTCCA 689
QY 929 taagggaagaactaaagagaacgcgataccattttctggc 967
|||||

Db 690 TAACGGGAACAACATAAG-GGACAGATACCAATATTCTGGC 727

RESULT 12

LOCUS W21436 412 bp mRNA EST 20-AUG-1996
DEFINITION zB52a06.r1 Soares_fetal_lung_NbH19W Homo sapiens cDNA clone
IMAGE:307186 5', mRNA sequence.
W21436
ACCESSION W21436
VERSION W21436.1 GI:1298078
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 412)
AUTHORS Hillier L., Lennon G., Becker M., Bonaldo M.F., Chiapelli B.,
Chissoe S., Dietrich N., Dubuque T., Favello A., Gish W., Hawkins
M., Hultman M., Kucaba T., Lacy M., Le M., Le N., Mardis E., Moore
B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T.,
Schellenberg K., Soares M.B., Tan F., Thierry-Mieg J., Trevaskis E.,
Underwood K., Wohldmann P., Waterston R., Wilson R. and Marra M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2452 Std Error: 0.00
Seq primer: mob.REGA+ET

FEATURES

Location/Qualifiers
1..412
/organism="Homo sapiens"
/db_xref="GDB:1250598"
/db_xref="taxon:9606"
/clone="IMAGE:307186"
/clone_lib="Soares_fetal_lung_NbH19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCCCAATTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbH19W."
BASE COUNT 117 a 78 c 92 g 125 t
ORIGIN

Query Match 23.3%; Score 411; DB 190; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.9e-101;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 459 catatgcagaatgatttttagaagttgaggaacccagtgatggaactatattagggcg 518
|||||
Db 1 CATATGCAAGTATGATTTTGTAGAAGTTGAGAACCCAGTGTGATGATGATGATGATGATG 60
QY 519 ctggtgtggtctggtactgtaccaggaacacagatttcttaagaaataaattagat 578
|||||
Db 61 CTGGTGTGGTCTGGTACTGTACAGGAGAAACAGATTTCTAAAGGAAATTAATAGGAT 120
QY 579 aagattgtatctgatgaattttctctgaaccagggttctctgcatccactacaacat 638

Db 121 AAGATTGTATCGATGAATATTTCCCTTCTAACACAGGTTCTGTCATCCACTTACAAACAT 180
QY 639 tgtcatgccacaattcacagaagctgtgagctcttcagtgctaccccccttcagctttgcc 698
Db 181 TGTGATGCCAATTCACAGAAGCTGTGAGTCTCTCAGTGTGTCACCCCTTCAGCTTTGCC 240
QY 699 actgacctgcttaataatgctataactgcctttagctacttgggaagacctatttcgata 758
Db 241 ACTGGACCTGCTTAATATGCTATAACTGCCCTTAGTACTCTGGAGAGACTTATTTCGATA 300
QY 759 tcttgaccagagagatgacagcttgagcttagaagatctatatagggccaaacttggcaact 818
Db 301 TCTTGAACAGAGATGCGACTGTGACTTAGAAGATCTATATAGGCCAACTTGGCAACT 360
QY 819 tcttgcaagctcttctgttttgggaagaaatccagagtggtggtatctgaa 869
Db 361 TCTTGGCAAGGCTTTGTGTTTGGAGAAATCCAGAGTGGTGGATCTGAA 411

RESULT 13
BF151355 564 bp mRNA EST 29-DEC-2000
LOCUS uz15b12.y1 NCI-CGAP_Mam5 Mus musculus cDNA clone IMAGE:3669119 5'
DEFINITION similar to TR:Q90Y71 Q90Y71 FALLOTEIN. ; mRNA sequence.
ACCESSION BF151355
VERSION BF151355.1 GI:11032750
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 564)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1429887
Seq primer: -40RP from Gibco
High quality sequence stop: 436.
Location/Qualifiers
1. 564
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:3669119"
/clone_lib="NCI-CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 157 a 130 c 140 g 137 t
ORIGIN

Query Match 23.1%; Score 407.8; DB 145; Length 564;
Best Local Similarity 82.8%; Pred. No. 1.5e-100;
Matches 466; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 619 tctgtccactacacaattgtcatcgcacaaatttcacagaagctgtgagctcttcagtg 678
Db 1 TTCTGTCATCCACTACAGTATTATTCATGCCACAAGTCACAGAAACACACGAGTCTTCGCTG 60
QY 679 ctaccccttcagcttggcaactggaacctgcttaataatgctataactgcttagtacc 738
Db 61 TTGCCCCCTTTCATCTTTGTGCTATTGGACCTGCTCAACATGCTGCTACCTTCAGTACC 120
QY 739 ttggaagaccttattcgatatcttgaacagagagatggcagtggaacttgaagatata 798
Db 121 TTGGAAGAGCTGATTCGGTACCTAGAGCCAGATCGATGGCAGGTGGACTTGGACAGCCTC 180
QY 799 tataggccaacttggcaactcttggcaagcttttgggttttgggaagaaaaatccagatg 858
Db 181 TACAAGCCACATGGCAGCTTTTGGCAAGCTTTCTGTATGGAAAAAAGCAAAATG 240
QY 859 gtgagatcgaaaccttcaacagagaggttaagattatatacagctgcacaccttgtaactc 918
Db 241 GTGAATCTGAATCTCTCAAGGAAGAGGTAAACTCTACAGCTGCACACCCCGGAACCTC 300
QY 919 tcaagtccataaaggaagaactaaagaaacgaacacacatttcttgcaggttgcctc 978
Db 301 TCAGTGTCATACACGGGAAGAGCTAAAGAGGACAGATACCATATTCTGGCCAGGTGTCTC 360
QY 979 ctggttaaacgctgtgtgggaactgtgcctgttgcctcccaattgcaatgaatgtcaa 1038
Db 361 CTGGTCAAGCGCTGTGGAGGAATTTGCCCTGTTGCTCCATTAATGCAATGAATGTGAC 420
QY 1039 tgtgtcccaagcaaaagtactaaaaataccacagaggttccttcagttgagaccacaaagacc 1098
Db 421 TGTGTCCACGTAAGTTACAAAAAAGTACCATGAGGTCTTCAGTTGAGACCAAAAACT 480
QY 1099 ggtgtcagggattgcaaaaatacactcacgacgtggccctggagcaccatgaggagtgt 1158
Db 481 GGAGTCAAGGGATGTCATAAAGTACTCAATGATGTGCTCTTGAACACACCGAGGAATGT 540
QY 1159 gactgtgtgtgcagaggagcac 1181
Db 541 GACTGGGTGTGAGAGGGAACGC 563

RESULT 14
BG112388 983 bp mRNA EST 30-JAN-2001
LOCUS 602281957F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369430 5',
DEFINITION mRNA sequence.
ACCESSION BG112388
VERSION BG112388.1 GI:12605894
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 983)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM10025 row: a column: 15
High quality sequence stop: 397.
Location/Qualifiers
1. 983
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4369430"

FEATURES
source

Job time: 4474 sec

7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29., 2001, 17:05:08 ; Search time 22.8 seconds

(without alignments)
917.337 Million cell updates/sec

Title: US-09-457-066-2

Perfect score: 1858

Sequence: 1 MSLLFGLLLLSALAGORQGT.....DVALEHHECDVCVRGSGTG 345

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_0601.*

```

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1858	100.0	345	20	Human VEGF-E prote
2	1858	100.0	345	20	Human PRO200 prote
3	1858	100.0	345	20	Human vascular end
4	1858	100.0	345	21	Human zveg3, SEQ
5	1858	100.0	345	21	Human platelet-der
6	1858	100.0	345	21	Human PRO200 (UN01
7	1858	100.0	345	21	Human RACE generat
8	1858	100.0	345	21	Human VEGF-X prote
9	1858	100.0	345	21	Human VEGF-X prote
10	1858	100.0	345	21	Human 990126veg p
11	1858	100.0	345	21	Human VEGF-X prote

12	1858	100.0	345	21	AA19578
13	1858	100.0	345	21	AA19578
14	1858	100.0	345	21	AA19578
15	1858	100.0	345	21	AA19578
16	1858	100.0	345	21	AA19578
17	1858	100.0	345	21	AA19578
18	1858	100.0	345	21	AA19578
19	1858	100.0	345	21	AA19578
20	1858	100.0	345	21	AA19578
21	1858	100.0	345	21	AA19578
22	1851	99.6	345	21	AA19578
23	1850	99.6	345	21	AA19578
24	1850	99.6	345	21	AA19578
25	1850	99.6	345	21	AA19578
26	1745	93.9	345	21	AA19578
27	1745	93.9	345	21	AA19578
28	1728	93.0	318	21	AA19578
29	1671.5	90.0	339	21	AA19578
30	1667	89.7	345	21	AA19578
31	1667	89.7	345	21	AA19578
32	1667	89.7	345	21	AA19578
33	1450.5	78.1	282	21	AA19578
34	1316	70.8	297	21	AA19578
35	814	43.8	167	21	AA19578
36	798	42.9	168	21	AA19578
37	753	40.5	180	21	AA19578
38	742.5	40.0	370	21	AA19578
39	742.5	40.0	370	21	AA19578
40	742.5	40.0	370	21	AA19578
41	742.5	40.0	370	21	AA19578
42	739.5	39.8	322	21	AA19578
43	737.5	39.7	370	21	AA19578
44	735.5	39.6	370	21	AA19578
45	646	34.8	149	21	AA19578

ALIGNMENTS

RESULT 1

AA19578

ID AA19578 standard; Protein; 345 AA.

AC AA19578;

DT 11-JAN-2000 (first entry)

DE Human VEGF-E protein.

DE VEGF-E; human; vascular endothelial cell growth factor; wound repair;

KW treatment; cardiovascular disorder; endothelial disorder; therapy;

KW tissue generation; regeneration; cardiac hypertrophy; cancer; detection;

KW angiogenic disorder; age-related macular degeneration; vascular disease;

KW neovascularization; tumor; gene mapping.

OS Homo sapiens.

PN WO9947677-A2.

PD 23-SEP-1999.

PF 10-MAR-1999; 99WO-US05190.

PR 17-MAR-1998; 98US-0040220.

PR 02-NOV-1998; 98US-0184216.

PA (GETH) GENENTECH INC.

PI Ferrara N, Kuo SS;

PI WPI; 1999-580306/49.

DR N-PSDB; AA23691.

XX

PT New growth factor polypeptide useful for treating cardiovascular or
PT endothelial disorders, e.g. cardiac hypertrophy -
XX Claim 1; Fig 2; 122pp; English.
XX
CC This invention describes the isolation of a novel human vascular
CC endothelial cell growth factor-E (VEGF-E) polypeptide which has
CC tranquilizer, vulnerary and cardiant activity. VEGF-E can be administered
CC therapeutically, especially by expressing encoding polynucleotides, to
CC treat cardiovascular or endothelial disorders in mammals, especially
CC humans. It is useful in wound repair and tissue generation and
CC regeneration, and may especially be used to treat cardiac hypertrophy
CC It can be combined with a carrier in pharmaceutical compositions, which
CC can be administered to treat disorders as above. VEGF-E can be used to
CC screen for antagonists and agonists, and the antagonists administered to
CC treat angiogenic disorders in mammals (especially humans) e.g. cancer or
CC age-related macular degeneration. It can be used to generate antibodies,
CC useful therapeutically as antagonists, as above. The antibodies are also
CC useful to detect VEGF-E polypeptide, especially to diagnose
CC cardiovascular, endothelial or angiogenic disorders in mammals (e.g.
CC vascular disease, or neovascularization associated with tumor formation),
CC by contacting the antibody with a tissue sample and detecting formation
CC of an antibody-VEGF-E polypeptide complex. Polynucleotides encoding
CC VEGF-E can be used to diagnose cardiovascular and endothelial disorders
CC in mammals, by detecting abnormally high or low VEGF-E gene expression in
CC tissue samples. They can also be used to diagnose a disease or
CC susceptibility to a disease related to a mutated form of VEGF-E (e.g. a
CC cardiovascular, endothelial or angiogenic disorder such as a tumor), by
CC detecting a mutation in the VEGF-E encoding sequence isolated from a
CC sample. They may also be used to produce probes useful to detect related
CC sequences or for gene mapping. This sequence represents the human VEGF-E
CC protein described in the method of the invention.
XX Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 20; Length 345;
Best Local Similarity 100.0%; Pred. NO. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFLGILLTSALAGRGQTAEENLSKQFSSNKQNGVDPQHERIIVTNGSTHS 60
Db 1 msflgilltsalagrqtgaenlskqfssnkqngvdpqheriivtngsths 60
QY 61 PRFHTYPRNTVLWRLVAEENWVQLTDERFGLDEPDIDCKYDFVEEPPSDGTIL 120
Db 61 prfhtyprntvlwrlvaeenwvqltderfglepdidckdfveeppsdtgil 120
QY 121 GRWCGSTVPKQISKGNQIRFVSDYFPPSEPGFCIHYNVMPQFTFVSPVLPSPA 180
Db 121 grwcgsgtvpkqiskgnqirfvsdyfppsepgfcihynvmpqftfveavspvlpssa 180
QY 181 LPDLNNAITAFSTLEDLTRYLEPERWQDLEDLYRPTWQLLGAFFGKSRVDNLN 240
Db 181 lpdlannaatstledltryleperwqldledlyrptwqlqkafvgrksrvvdlnl 240
QY 241 LTEEVRLYSTPRNFVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCQCQVPSK 300
Db 241 lteevrlystprnfvsireelkrtdtifwpgcllvkrcgncacclhncqcqvpsk 300
QY 301 VTKKYHEVLQRPKTVGRGLHKSITDVALEHHECDVCVRGSGTG 345
Db 301 vtkeyhavlqrpktvgrglhksitdvalhheecdcvcrgstg 345

RESULT 2

AA41766

ID AA41766 standard; Protein: 345 AA.

XX AA41766;

AC AA41766;

XX 07-DEC-1999 (first entry)

DT

XX

DE Human PRO200 protein sequence.
XX Human; PRO: EST; expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW secreted protein; transmembrane protein.
OS Homo sapiens.
XX WO9946281-A2.
PN 16-SEP-1999.
XX 08-MAR-1999; 99WO-US05028.
XX 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 09-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.

Db	1	mslfigllltsalagrqrqtqæälnlskfkfsnkeqngvqdpqcheriitvstngsihs	60
QY	61	PREPHTYPRNTVLVWLVAEENVMWIOQLTFDERFLEDPEDDICKYDFVEVEEPSDGTIL	120
Db	61	prfphyrprntvlvwlvaeenvvqitfderfgledpeddickdyfveeepsdgtil	120
QY	121	GRWGSSTVPGKOISKGNQIRIRFVSDFVFPSEPGFCITHYNTVMPOQTFEAVSPSVLPESA	180
Db	121	grwgs9stvpgkqiskgnqirirfvsd5yfsepgfcibhnlvmppqfleaavspsvlpesa	180
QY	181	LPDLNNAITAFSTLEDILRYLEPERWQLDLEDILYRPTWQLLGKAFVGRKSRVVDLNL	240
Db	181	lpdl1nnaitafstledilryleperwqldledilrptwqllgkafvgrksrvvdlnl	240
QY	241	LTEEVRLYSTPRNFVSVIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK	300
Db	241	lteevrlystprnfsvsireelkrtdtifwpgcllvrcgncacclhncnecqvpsk	300
QY	301	VTKKYHEVLQRLPKTGVRLGHLKSLTDVALEHHEECDCVCRGSTGG	345
Db	301	vtkkyhevlqrlpaktvgrlgklsldvalenhheecdcvcrstgg	345

RESULT	4	
AAAB48657		
ID	AAAB48657	standard; Protein; 345 AA.
XX		
AC	AAAB48657;	
XX		
DT	09-MAR-2001	(first entry)
XX		
XX	Human zvegf3	SEQ ID NO:33.
DE		
XX		
XX	Human; zvegf3; zvegf4 fusion; growth factor homologue; VEGF/PDGF family;	
KW	CUB domain; PDGF-like activity; mitogenic; osteogenic;	
KW	neovascularisation; tissue repair; proliferation; differentiation;	
KW	liver damage; neuroregenerative; Alzheimer's disease; multiple sclerosis;	
KW	periodontal disease; bone fracture; wound healing; vulnery; ischaemia;	
KW	immunomodulation; hepatic.	
XX		
OS	Homo sapiens.	
XX		
XX	WO2000066736-A1.	
PN		
XX		
XX	09-NOV-2000.	
PD		
XX		
XX	03-MAY-2000; 2000WO-US40047.	
PF		
XX		
XX	03-MAY-1999; 99US-0304216.	
PR		
XX	10-NOV-1999; 99US-0164463.	
PR		
XX	04-FEB-2000; 2000US-0180169.	
XX		
PA	(ZYMO) ZYMOGENETICS INC.	
XX		
PI	Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;	
XX		
XX	WPI; 2000-687541/67.	
DR		
DR	N-PSDB; AAC81582.	
XX		
PT	Growth factor homologs and the nucleic acids that encode them, useful	
PT	e.g. for treating liver damage, ischemia, multiple sclerosis and	
PT	Alzheimer's disease -	
XX		
XX	Claim 48; Page 125-126; 143pp; English.	
PS		
XX		
XX	The invention relates to the human growth factor homologue zvegf4	
CC	(AAAB48653), and nucleic acids encoding it (AAC81555). Zvegf4 is a member	
CC	of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial	
CC	growth factor) family. Zvegf4 has a growth factor domain (AAAB48654)	
CC	characterised by a PDGF cysteine knot structure, and a CUB domain	
CC	(AAAB48655) which has a beta barrel structure. Zvegf4 has PDGF-like	
CC	activity, having mitogenic activity on fibroblasts, vascular smooth	
CC	muscle cells and pericytes, and has also been shown to stimulate bone	

growth. The invention also relates to fusion proteins comprising human zvegf4 or fragments thereof, particularly human zvegf4/human zvegf3 fusions; expression constructs and host cells comprising human zvegf4 nucleic acids; the recombinant expression of human zvegf4; an antibody which binds to human zvegf4 or a fragment thereof; a method of activating a cell-surface PDGF receptor using a zvegf4-derived polypeptide; a method of modulating the proliferation, differentiation, migration or metabolism of bone cells, comprising exposing bone cells to zvegf4-derived polypeptides; and a method of detecting a genetic abnormality in the zvegf4 gene of a patient. zvegf4 proteins and derived fragments may be used to stimulate tissue development or repair, or cellular differentiation or proliferation. They are particularly used for the treatment or repair of liver damage, and may also be used to modulate neurite growth (e.g., in the treatment of Alzheimer's disease or multiple sclerosis). Due to their osteogenic activity, they may be used in the treatment of periodontal disease and fractures. They may also be used to enhance expansion and mobilisation of haematopoietic stem cells and endothelial precursor stem cells, which may be useful in the treatment of ischaemia, in wound healing, and in the modulation of the immune system. The present sequence represents human zvegf3.

Sequence 345 AA;

Query Match	100.0%	Score 1858;	DB 21;	Length 345;
Best Local Similarity	100.0%;	Pred. No. 6e-180;		
Matches 345;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSLFGLLLTALAGORGTGAESLSSKFFQSSNKQONGVQDPOHERITITVSTNGSIHS	60	
Db	1	mslfgllltsalaggrgtgaesnlsskfgtssnkeqngvqdpqnerlittvstngsihs	60	
QY	61	PRFPHTYPRNTVLVWRLVAEENWMIQTLTDFERFQLEDPEDDICKYDFVEVEEPSDGTIL	120	
Db	61	prfphtpyntvlvwrLvaveenwigtldferfgleddedickdfveveepsdgttl	120	
QY	121	GRWCGSTVPQKQISKGNQIRIRFVSDYFSPSEPGFCHYNIIVMPQFTAEAVSPSLVPSPA	180	
Db	121	grwcgstvpqkqiskgnqirirfvsdyfpsepgfchyniivmpqfteavspslvpspa	180	
QY	181	LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLIRPTWQLLGKAFVFGKRSRVVDLNL	240	
Db	181	lpdlnnaitafstledliryleperwqldledlirptwqllgkafvfgkrsvvdlnl	240	
QY	241	LTEEVRLYSCTPRNFVSIREELKRTDIFWPGCILLVRKCGNCACCLHNCMECCQVPSK	300	
Db	241	lteevrlyscptrnfvsireelkrtdifwpgcillvrkcgncacclhncmecqvpksk	300	
QY	301	VTKKYHEVLQRPKTGVGRGLHKSITDVALEHHHEECDCVCRGSTGG	345	
Db	301	vtkkyhevlqrpktgvgrglhksitdvaaleheecdcvcrastgg	345	

RESULT	5	
AAB24250		
ID	AAB24250	standard; protein; 345 AA.
XX	XX	
AC	AAB24250;	
XX	XX	
DT	08-FEB-2001	(first entry)
XX	XX	
DE	Human	platelet-derived growth factor related protein LP8.
XX	XX	
KW	Human;	platelet derived growth factor related protein; LP8; VEGFh;
KW	vascular endothelial	growth factor h; tissue regeneration; vulnerary;
KW	atherosclerosis;	PDGF-related protein; antiarteriosclerotic.
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	WO200059940-A2.	
XX	XX	
PD	12-OCT-2000.	
XX	XX	

PF 24-MAR-2000; 2000WO-US06427.
XX
PR 06-APR-1999; 99US-0127913.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Hammond LJ, Na S;
XX
DR WPI; 2000-664991/64.
DR N-PSDB; AAC64426.
XX
PT Enhancing tissue growth and promoting wound healing by administering
PT platelet-derived growth factor related protein, LP8 or its analog and
PT treating atherosclerosis by administering LP8 antagonist
XX
PS Claim 4; Page 63-64; 64pp; English.
XX
CC The present invention describes a method for enhancing tissue growth,
CC promoting wound healing or stimulating smooth muscle growth by
CC administering a platelet-derived growth factor (PDGF) related protein,
CC designated LP8 or its analogue. Also described is a method of slowing
CC the progress of atherosclerosis or treating atherosclerosis comprising
CC the administration of an LP8 antagonist. The method is useful for
CC enhancing tissue growth, promoting wound healing and stimulating smooth
CC muscle growth. Antagonists of LP8 are useful for treating
CC atherosclerosis. The present sequence represents human LP8, which is
CC also called VEGFh.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGORQGTQAESNLSSKQFSSNKQNGVQDPQHERITVTSTNGSIHS 60
DB 1 mslfglilltsalagrgtqaesnlsskfssnkqngvqdpqheriitvtstngsihs 60

QY 61 PRPHPTYPRNTVLVRLVAEENWVLIQTFDERFGLDEPDICKYDFVEVEEPSDGTIL 120
DB 61 prfphptyprntvlvrlvaeenwvliqtfderfgledpeddickdyfveveepsdgtil 120

QY 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPVLPPSA 180
DB 121 grwcsgstvpkgqiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlppsa 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWOLDLEDLYRPTWOLLGKAFVGRKSRVVDNL 240
DB 181 lpldlnnaitafstledliryleperwoldledlyrptwollgkafvgrksrvvldnl 240

QY 241 LTEEVRVLSCTPRNFSVSIREEKRTDITFWPGCLLVKRCGNACCLHNCNECQVPSK 300
DB 241 lteevrlyscprnfsvsiireekrtditfwpgcllvkrcgnacclhncnecqvpsk 300

QY 301 VTKKYHEVILQRPKGVRLGKSLTDVALEHHHEDCVCRGSTGG 345
DB 301 vtkeyhevilqrpkgtvrglksltdvalehheedcvcrgstgg 345

RESULT 6
AAB44322
ID AAB44322 standard; Protein; 345 AA.
XX
AC AAB44322;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO200 (UNQ174) protein sequence SEQ ID NO:488.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytosstatic;
KW expressed sequence tag; detection; cancer.
XX

OS Homo sapiens.
XX
PN WO200053756-A2.
XX
PD 14-SEP-2000.
XX
PF 18-FEB-2000; 2000WO-US04341.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX
DR WPI; 2000-611443/58.
DR N-PSDB; AAC78582.
XX
PT Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities -
XX
PS Claim 12; Fig 207; 636pp; English.
XX
CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting
CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78997 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGORQGTQAESNLSSKQFSSNKQNGVQDPQHERITVTSTNGSIHS 60
DB 1 mslfglilltsalagrgtqaesnlsskfssnkqngvqdpqheriitvtstngsihs 60

QY 61 PRPHPTYPRNTVLVRLVAEENWVLIQTFDERFGLDEPDICKYDFVEVEEPSDGTIL 120
DB 61 prfphptyprntvlvrlvaeenwvliqtfderfgledpeddickdyfveveepsdgtil 120

QY 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPVLPPSA 180
DB 121 grwcsgstvpkgqiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlppsa 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWOLDLEDLYRPTWOLLGKAFVGRKSRVVDNL 240
DB 181 lpldlnnaitafstledliryleperwoldledlyrptwollgkafvgrksrvvldnl 240

QY 241 LTEEVRVLSCTPRNFSVSIREEKRTDITFWPGCLLVKRCGNACCLHNCNECQVPSK 300
DB 241 lteevrlyscprnfsvsiireekrtditfwpgcllvkrcgnacclhncnecqvpsk 300

QY 301 VTKKYHEVILQRPKGVRLGKSLTDVALEHHHEDCVCRGSTGG 345
DB 301 vtkeyhevilqrpkgtvrglksltdvalehheedcvcrgstgg 345

RESULT 6
AAB44322
ID AAB44322 standard; Protein; 345 AA.
XX
AC AAB44322;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO200 (UNQ174) protein sequence SEQ ID NO:488.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytosstatic;
KW expressed sequence tag; detection; cancer.
XX

Db 121 grwcgsvtpgkqskgnqirirfsvdeyfpsepgfcihynivmpqfteavspvlpssa 180

Qy 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240

Db 181 lpldllnnaitafstledliryleperwqldledlyrptwqlgkafvgrksrvvdnl 240

Qy 241 LTEEVRVLYSCTPRNFSVIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300

Db 241 lteevrlyscptrnfsvsireelkrttdtlfwpgcllvkrcgncacclhncnecqvpksk 300

Qy 301 VTKKYHEVLQRPKTGVGRGLHKSITDVALEHHEECDCVCRGSTGG 345

Db 301 vtkeyhevlqirphtgvrghlksitdvalheheecdcvcrgstgg 345

RESULT 7

AAB10633

ID AAB10633 standard; Protein; 345 AA.

XX

AC AAB10633;

XX

DT 19-JAN-2001 (first entry)

XX

DE Human RACE generated VEGF-X protein.

XX

KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic; antiarthritis; antipsoriatic; antidiabetic; treatment; angiogenesis regulator; vascularization regulator; cancer; psoriasis; rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair; tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore; venous sore; diabetic ulcer; burns; skin graft growth.

XX

OS Homo sapiens.

XX

PN WO200037641-A2.

XX

PD 29-JUN-2000.

XX

PF 21-DEC-1999; 99WO-US30503.

XX

PR 22-DEC-1998; 98GB-0028377.

PR 18-MAR-1999; 99US-0124967.

PR 08-NOV-1999; 99US-0164131.

XX

PA (JANC) JANSSEN PHARM NV.

XX

PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gosiewska A;

PI Dhanaraj SN, Xu J;

XX

XX WPI: 2000-442669/38.

DR N-PSDB; AAA71951.

XX

PT New vascular endothelial growth factor protein, useful for treating or preventing diseases associated with inappropriate angiogenesis activity such as cancer, rheumatoid arthritis, psoriasis and wounds -

XX

PS Disclosure; Fig 6; 127pp; English.

XX

CC This invention describes a novel vascular endothelial growth factor-X (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has vulnery, cytostatic, antiarthritis, antipsoriatic, antidiabetic and angiogenesis regulator activity and acts as an angiogenesis and vascularization regulator. An antisense molecule of the invention is useful for treating or preventing cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy by inhibiting angiogenic activity or inappropriate vascularization including formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject. The products of the invention are useful for preparing medicaments for treating wounds such as dermal ulcers, pressure sores, venous sores, diabetic ulcers and burns and to promote skin graft growth, tissue repair, proliferation of new blood vessels, tissue regeneration and organ repair by promoting angiogenic activity or vascularization. This sequence represents the RACE generated human VEGF-X

CC protein described in the method of the invention.

XX

SO Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;

Best Local Similarity 100.0%; Pred. No. 6e-180;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLFGLLLTSALAGRQGTQAESNLSSKRFQFSSNKQNGVQDPQHERITITVSTNGSIHS 60

Db 1 mslflllltsalagrqtgaesnlsskrfssnkqngvqdpqheriitvstngsihs 60

Qy 61 PRFPHYTPRNTVLRVLRVAEENWVQLTDFDERFGLDEPDDICKYDFVEVEEPSDGTIL 120

Db 61 prfptyprntvrvlrvaeevwnvqltderfgleddedickdyfveveepsdgtill 120

Qy 121 GRWCGSVTPGKQISKGNQIRIRFVSDVEYFPSEPGFCIHYNIVMPQFTAVSPVLPSSA 180

Db 121 grwcgsvtpgkqiskgnqirirfsvdeyfpsepgfcihynivmpqfteavspvlpssa 180

Qy 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240

Db 181 lpldllnnaitafstledliryleperwqldledlyrptwqlgkafvgrksrvvdnl 240

Qy 241 LTEEVRVLYSCTPRNFSVIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300

Db 241 lteevrlyscptrnfsvsireelkrttdtlfwpgcllvkrcgncacclhncnecqvpksk 300

Qy 301 VTKKYHEVLQRPKTGVGRGLHKSITDVALEHHEECDCVCRGSTGG 345

Db 301 vtkeyhevlqirphtgvrghlksitdvalheheecdcvcrgstgg 345

RESULT 8

AAB10635

ID AAB10635 standard; Protein; 345 AA.

XX

AC AAB10635;

XX

DT 19-JAN-2001 (first entry)

XX

DE Human VEGF-X protein #1 isolated from clones 4 and 7.

XX

KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic; antiarthritis; antipsoriatic; antidiabetic; treatment; angiogenesis regulator; vascularization regulator; cancer; psoriasis; rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair; tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore; venous sore; diabetic ulcer; burns; skin graft growth.

XX

OS Homo sapiens.

XX

PN WO200037641-A2.

XX

PD 29-JUN-2000.

XX

PF 21-DEC-1999; 99WO-US30503.

XX

PR 22-DEC-1998; 98GB-0028377.

PR 18-MAR-1999; 99US-0124967.

PR 08-NOV-1999; 99US-0164131.

XX

PA (JANC) JANSSEN PHARM NV.

XX

PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gosiewska A;

PI Dhanaraj SN, Xu J;

XX

XX WPI: 2000-442669/38.

DR N-PSDB; AAA71955.

XX

PT New vascular endothelial growth factor protein, useful for treating or preventing diseases associated with inappropriate angiogenesis activity

PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX Disclosure; Fig 9; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-x
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers,
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents the human VEGF-X protein
CC isolated from clones 4 and 7 described in the method of the invention.
XX
XX Sequence 345 AA;
SQ

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLFGLLLTSALAGORGTQAESNLSSKFQSSNKEQNGVQDPQHERITVTSTNGSIHS 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 mslfgl|||ltsalagrggtqaesnlsskfqsnnkeqngvqdpqheritvtstngsihs 60

Qy 61 PRFPHYPRNTVLVRLVAVENWVQLTFDERFGLDEPDDICKYDFVEEPSDGTIL 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 prfphyprntvlvrlvaveenwvqltfderrfgldedpddickdyfveepsdgtil 120

Qy 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEFPSPGFCIHYNVMPQFTEAVSPVLPSPA 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 grwcsgstvpkgqiskgnqirirfvsdeyfpsepgfcihynvmqfteavspvlpssa 180

Qy 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLRYPTWQLLGKAFVGRKSRVVDLNL 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 lpdllnnaifstledliryleperwqldledlryptwqllgkafvgrksrvvdlnl 240

Qy 241 LTEEVRLYSCTPRNFSVSIREELKRTDIFWPGCLLVKRCGNCACCLHNCNCCQVPSK 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 lteevrlyscprnfsvsireelkrtidfwpgc|||lvkrcgncacclhncnccqvpsk 300

Qy 301 VTKKYHEVLQRLPKTGVRGLHKS LTDVALEHHEEDCDVCVRGSGTG 345
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 vtkkyhevlqrlpktgvrghksltdvalehheecdvcvrgstgg 345

RESULT 9
AAB10644
ID AAB10644 standard; Protein; 345 AA.
XX
AC AAB10644;
XX
XX 19-JAN-2001 (first entry)
DT
XX Human VEGF-X protein #4.
DE
XX VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth.
XX
XX Homo sapiens.
XX
XX WO200037641-A2.
PN

XX 29-JUN-2000.
XX
XX 21-DEC-1999; 99WO-US30503.
XX
XX 22-DEC-1998; 98GB-0028377.
PR 18-MAR-1999; 99US-0124967.
PR 08-NOV-1999; 99US-0164131.
XX (JANC) JANSSEN PHARM NV.
XX
XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
PI Dhanaraj SN, Xu J;
XX
XX WPI: 2000-442669/38.
DR N-PSDB; AAA71990.
XX
XX New vascular endothelial growth factor protein, useful for treating or
PT preventing diseases associated with inappropriate angiogenesis activity
PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Disclosure; Fig 30B; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-x
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers,
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents a human VEGF-X protein
CC described in the method of the invention.
XX
XX Sequence 345 AA;
SQ

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLFGLLLTSALAGORGTQAESNLSSKFQSSNKEQNGVQDPQHERITVTSTNGSIHS 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 mslfgl|||ltsalagrggtqaesnlsskfqsnnkeqngvqdpqheritvtstngsihs 60

Qy 61 PRFPHYPRNTVLVRLVAVENWVQLTFDERFGLDEPDDICKYDFVEEPSDGTIL 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 prfphyprntvlvrlvaveenwvqltfderrfgldedpddickdyfveepsdgtil 120

Qy 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEFPSPGFCIHYNVMPQFTEAVSPVLPSPA 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 grwcsgstvpkgqiskgnqirirfvsdeyfpsepgfcihynvmqfteavspvlpssa 180

Qy 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLRYPTWQLLGKAFVGRKSRVVDLNL 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 lpdllnnaifstledliryleperwqldledlryptwqllgkafvgrksrvvdlnl 240

Qy 241 LTEEVRLYSCTPRNFSVSIREELKRTDIFWPGCLLVKRCGNCACCLHNCNCCQVPSK 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 lteevrlyscprnfsvsireelkrtidfwpgc|||lvkrcgncacclhncnccqvpsk 300

Qy 301 VTKKYHEVLQRLPKTGVRGLHKS LTDVALEHHEEDCDVCVRGSGTG 345
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 vtkkyhevlqrlpktgvrghksltdvalehheecdvcvrgstgg 345

RESULT 10

AAB10650	121	GRWCGSTVFGKQISKGNQIRIRFVSDEYFPSPGFCIHYNVMPQFTFVSPSVLPSPA	180
ID	AAB10650	standard; Protein; 345 AA.	
XX	AC	AAB10650;	
XX	DT	19-JAN-2001 (first entry)	
XX	DE	Human 990126vegx protein.	
XX	KW	VEGF-X; vascular endothelial growth factor; human; vulnery; cytotstatic;	
XX	KW	antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;	
XX	KW	angiogenesis regulator; vascularization regulator; cancer; psoriasis;	
XX	KW	rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;	
XX	KW	tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;	
XX	KW	venous sore; diabetic ulcer; burns; skin graft growth.	
XX	OS	Homo sapiens.	
XX	PN	WO200037641-A2.	
XX	PD	29-JUN-2000.	
XX	PF	21-DEC-1999; 99WO-US30503.	
XX	PR	22-DEC-1998; 98GB-0028377.	
XX	PR	18-MAR-1999; 99US-0124967.	
XX	PR	08-NOV-1999; 99US-0164131.	
XX	PA	(JANC) JANSSEN PHARM NV.	
XX	PI	Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;	
XX	PI	Dhanaraj SN, Xu J;	
XX	PD	WPI; 2000-442669/38.	
XX	PT	New vascular endothelial growth factor protein, useful for treating or	
XX	PT	preventing diseases associated with inappropriate angiogenesis activity	
XX	PT	such as cancer, rheumatoid arthritis, psoriasis and wounds -	
XX	PS	Disclosure; Fig 11; 127pp; English.	
XX	CC	This invention describes a novel vascular endothelial growth factor-X	
XX	CC	(VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has	
XX	CC	vulnerary, cytotstatic, antirheumatic, antiarthritic, antipsoriatic and	
XX	CC	antidiabetic activity and acts as an angiogenesis and vascularization	
XX	CC	regulator. An antisense molecule of the invention is useful for treating	
XX	CC	or preventing cancer, rheumatoid arthritis, psoriasis and diabetic	
XX	CC	retinopathy by inhibiting angiogenic activity or inappropriate	
XX	CC	vascularization including formation and proliferation of new blood	
XX	CC	vessels, growth and development of tissues, tissue regeneration and organ	
XX	CC	and tissue repair in a subject. The products of the invention are useful	
XX	CC	for preparing medicaments for treating wounds such as dermal ulcers,	
XX	CC	pressure sores, venous sores, diabetic ulcers and burns and to promote	
XX	CC	skin graft growth, tissue repair, proliferation of new blood vessels,	
XX	CC	tissue regeneration and organ repair by promoting angiogenic activity or	
XX	CC	vascularization. This sequence represents the human 990126vegx protein	
XX	CC	used to illustrate the method of the invention.	
XX	XX	Sequence 345 AA;	
SQ			
Query Match	100.0%;	Score 1858;	DB 21; Length 345;
Best Local Similarity	100.0%;	Pred. No. 6e-180;	
Matches 345;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MSLFGILLTTSALAGRQGTQAEISLRSKFQSSNKEQNGVDPQHERITVTNGSIHS	60
DB	1	mslfgillttsalagrqtqaesnlsskfqsnskeqngvdpqheriitvtstngsihs	60
QY	61	PRPPTHTYNTVILWRLVAEENNVLTQTFDERFGLDEPDDICKYDFVEVEEPPSGTIL	120
DB	61	prfptyntvtilwrlvaeenvvltqtfderfgldepeddickdyfveveepsdgtill	120

CC vascularization. This sequence represents the human VEGF-X protein
CC described in the method of the invention.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGILLTSLALAGRQGTQAESNLSSKQFSSNKEQNGVQDPQOHERIITVSTNGSIHS 60
Db 1 msflgillttsalagrggtqaesnlsskqfssnkeqngvqdpqheriitvstngsihs 60

QY 61 PRPHPTVPRNTVLVRLVAVENWVQLTDFERFGLDEPDDICKYDFVEVEPSDGTIL 120
Db 61 prphptvprntvvlrvlavveenwvqltdfderfgleddedddickdyfvevepsdgtill 120

QY 121 GRWCGSGTVPKGQKISKGNQIRIRFVSDEYFPSPGFCIHYNIVMPQFTEAVSPSVLPSPA 180
Db 121 grwcgsgtvpkgqkiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspsvlppsa 180

QY 181 LPDLNNAITARSTLEDLIRYLEPERWQDLIEDLYRPTWOLLGKAPVGRKSRVVDNLN 240
Db 181 lpldlnnaitarstledliryleperwqdlledlyrptwollgkafvgrksrvvldnl 240

QY 241 LTEEVRLYSCTPRNFSVSTREELKRTDTIFWPGCLLVKRCGGNACCLHNCNECQVPSK 300
Db 241 lteevrlyscprnfsvsreelkrtdtifwpgcllvkrcggncacclhncnecqcvpsk 300

QY 301 VTKKYHEVLQRLPKTVGRGLHKSITDVALEHHEECDCVCRGSGTG 345
Db 301 vtkkyhevlqlrpkrtvgrglhksltadvalehheecdcvcrgstgg 345

RESULT 12
AAB19578
ID AAB19578 standard; Protein; 345 AA.
XX AAB19578;
AC AAB19578;
XX
DT 22-JAN-2001 (first entry)
DE Human PRO200 (vascular endothelial growth factor E).
XX
KW PRO200; vascular epithelial growth factor E; VEGF-E; human;
KW ocular disease; retinopathy; maculopathy; therapy;
KW retinitis pigmentosa; macular degeneration; retinal detachment;
KW retinal tear; macular hole; myopia; traumatic choriorretinopathy;
KW acute retinal necrosis syndrome; contusion; edema;
KW retinal vision occlusion; vascular disease; retinal vasculitis;
KW thrombocytopenic purpura; uveitis; retinal occlusion.
XX
OS Homo sapiens.
XX
FH
FT Key Location/Qualifiers
FT Peptide 1..14
FT Protein /label= Signal_peptide
FT 15..345
FT /label= Mature_Pro200
FT Modified-site 25..29
FT /note= "Asn is N-glycosylated"
FT Modified-site 55..59
FT /note= "Asn is N-glycosylated"
FT Modified-site 254..258
FT /note= "Asn is N-glycosylated"
FT Modified-site 15..21
FT /note= "N-myristoylation"
FT Modified-site 117..123
FT /note= "N-myristoylation"
FT Modified-site 127..133
FT /note= "N-myristoylation"
FT Modified-site 281..287

FT Modified-site /note= "N-myristoylation"
FT 282..288
FT /note= "N-myristoylation"
FT Modified-site 319..325
FT /note= "Amidation"
XX
XX WO2000053760-A2.
PN 14-SEP-2000.
XX
XX 10-MAR-2000; 2000WO-US063119.
XX
XX 12-MAR-1999; 99US-0123957.
XX (GETH) GENENTECH INC.
XX Ferrara N, Goddard A, Gurney AL, Hebert C, Henzel WJ, Kabakoff RC;
PI Klein RD, Kljavin IJ, Kuo SS, La Fleur M, Wood WI;
XX WPI: 2000-587437/55.
XX N-PSDB; AAA88515.
XX
XX Novel PRO polypeptides useful for preventing or rescuing retinal cells
XX from injury caused by ocular diseases such as retinitis pigmentosa,
XX retinopathy, retinal degenerative diseases, degenerative myopia,
XX uveitis -
XX
XX Claim 2; Fig 2; 140pp; English.
XX
XX The present sequence is that of human PRO200 or vascular
XX endothelial growth factor E (VEGF-E), as predicted from a cDNA
XX clone (see AAA88515) that was isolated from a glioma cell line G61
XX library using probes (see AAA88523-26) based on an expressed sequence
XX tag (see AAA88522) that showed homology to VEGF. PRO200 has a
XX predicted mol.wt. of 39,029 and a pI of about 6.06. A method for
XX producing PRO polypeptides, including PRO200, using a host cell
XX transformed with a vector comprising a PRO nucleic acid is claimed.
XX The invention relates to the use of PRO polypeptides to delay,
XX prevent or rescue retinal cells such as retinal neurons selected from
XX photoreceptors, retinal ganglion cells, displaced retinal ganglion
XX cells, amacrine cells, displaced amacrine cells, horizontal and
XX bipolar neurons, and supportive cells (including Mueller cells and
XX pigment epithelial cells) from injury and degradation. The retinal
XX cells are preferably photoreceptors and photoreceptor cell injury or
XX death is caused by retinal injury, light or environmental trauma or
XX by an ocular disease selected from retinitis pigmentosa, macular
XX degeneration, including age-related, retinal detachment, retinal
XX tears, retinopathy, retinal degenerative diseases, macular holes,
XX degenerative myopia, acute retinal necrosis syndrome, traumatic
XX choriorretinopathies or contusion such as Purtscher's retinopathy,
XX edema, ischemic conditions such as central or branch retinal vision
XX occlusion, collagen vascular diseases, thrombocytopenic purpura,
XX uveitis, retinal vasculitis and occlusion associated with Eales
XX disease and systemic lupus erythematosus (claimed).
XX
XX Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGILLTSLALAGRQGTQAESNLSSKQFSSNKEQNGVQDPQOHERIITVSTNGSIHS 60
Db 1 msflgillttsalagrggtqaesnlsskqfssnkeqngvqdpqheriitvstngsihs 60

QY 61 PRPHPTVPRNTVLVRLVAVENWVQLTDFERFGLDEPDDICKYDFVEVEPSDGTIL 120
Db 61 prphptvprntvvlrvlavveenwvqltdfderfgleddedddickdyfvevepsdgtill 120

QY 121 GRWCGSGTVPKGQKISKGNQIRIRFVSDEYFPSPGFCIHYNIVMPQFTEAVSPSVLPSPA 180
Db 121 grwcgsgtvpkgqkiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspsvlppsa 180

Db 241 lteevrlyscprnfsvsireelkrttdtfwpgcllvkrsgnccacclhmcncqcvpsk 300
Qy 301 VTKYHEVLQRLPKTGVLGHLKSLTDVALEHHEECDCVCRGSTGG 345
Db 301 vtkyhevqlrpkgtgvrghlksltdvalehheecdcvcrgstg 345

RESULT 14
AAB24412
ID AAB24412 standard; Protein; 345 AA.
XX
AC AAB24412;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human PRO713 protein sequence SEQ ID NO:137.
XX
KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;
KW cycostatic; gene therapy; vaccine.
OS Homo sapiens.
XX
PN WO200032221-A2.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US28313.
XX
PR 01-DEC-1998; 98WO-US25108.
PR 16-DEC-1998; 98US-0112850.
PR 12-JAN-1999; 99US-0115554.
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-OCT-1999; 99US-0162506.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
PI Watanabe CK, Williams PM, Wood WI;
XX
DR WPI; 2000-412154/35.
DR N-PSDB; AAA77621.

Nucleic acids encoding PRO polypeptides useful for preventing,
diagnosing and treating diagnosing a cardiovascular, endothelial or
angiogenic disorders in mammals -
XX
PS Claim 72; Fig 50; 315pp; English.

The present invention describes nucleic acids encoding PRO polypeptides
useful for preventing, diagnosing and treating diagnosing a
cardiovascular, endothelial or angiogenic disorder in mammals by
modulating cell proliferation, angiogenesis and cardiovascularisation,
and for identifying agonists and antagonists of these processes. The
nucleic acids and the proteins they encode may be used in the
prevention, treatment and diagnosis of diseases associated with
inappropriate PRO expression such as cardiovascular, endothelial or

CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAA77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.

XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSIFGLLLTSLALAGOROGTOAESNLSSKQFSSNKEONGVQDPOHERIITVSTNGSIHS 60
Db 1 msifglllltsalagrgtqaesnlsskfssnkeqngvqdpqheriitvstngsihs 60
Qy 61 PRFPHTYPRNTVLVRLVAEENWVQLTDFERFGLDEPDDEDDICKYDFVEVEPSPDGTIL 120
Db 61 prfphtyprntvlvrlvaeenwvqltderfgledpedddickdyfvevepsdgtiil 120
Qy 121 GRWCGSGTVPGKOISKGNQIRIRFVSDEVPFPGFCIHYNIVMPQTEAVSPSVLPSPA 180
Db 121 grwcsgstvpqkiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspsvlpssa 180
Qy 181 LPDLINNAITAFSTLEDLIRYLEPERWQDLEDLRYPTWQLLGKAFVGRKSRVVDNL 240
Db 181 lpdlinnaitafstledliryleperwqldledlryptwqllgkafvgrksrvvdnl 240
Qy 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHMCNECQCVPSK 300
Db 241 lteevrlyscprnfsvsireelkrttdtfwpgcllvkrsgnccacclhmcncqcvpsk 300
Qy 301 VTKYHEVLQRLPKTGVLGHLKSLTDVALEHHEECDCVCRGSTGG 345
Db 301 vtkyhevqlrpkgtgvrghlksltdvalehheecdcvcrgstg 345

RESULT 15

AAB01419
ID AAB01419 standard; Protein; 345 AA.

XX
AC AAB01419;

XX
DT 20-OCT-2000 (first entry)

XX
DE Human TANGO 128.

XX
KW TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
KW inflammatory bowel disease; septic shock; ulcerative colitis;
KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;
KW systemic lupus erythematosus; transgenic animal; diagnosis;
KW prognosis; prophylactic; therapeutic; human.

XX
OS Homo sapiens.

XX
PN WO200039284-A1.

XX
PD 06-JUL-2000.

XX
PF 23-DEC-1999; 99WO-US31025.

XX
PR 30-DEC-1998; 98US-0223546.

XX
PA (MILL-) MILLENNIUM PHARM INC.

XX
PI Holtzman DA;

XX
DR WPI; 2000-465743/40.

DR N-PSDB: AAA47452.

DR	N-PSDB; AAA47452.
XX	
PT	Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
PT	224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT	arthritis, psoriasis and autoimmune diseases

PS Claim 8; Fig 1; 209pp; English.

Nucleic acids encoding TANGO polypeptides are useful as modulating agents for regulating cellular processes like asthma, graft versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory bowel disease, septic shock, ulcerative colitis, Crohn's disease, chronic myelogenous leukemia, cancer, liver disease, Hodgkin's disease, osteoarthritis, Lyme's disease, cachexia and autoimmune diseases e.g. myasthenia gravis, autoimmune diabetes and systemic lupus erythematosus. The nucleic acids are also useful for producing transgenic animals and the TANGO polypeptides themselves. Partial TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in forensic biology, for diagnostic assays, prognostic assays, pharmacogenomics and for monitoring clinical trials. TANGO polypeptides are suitable for both prophylactic and therapeutic methods for treating a subject at risk of a disorder or having a disorder associated with aberrant TANGO expression. A wide range of cellular disorders can be treated.

Sequence 345 AA:

[illegible]

Search completed: June 29, 2001, 17:07:42
Job time: 154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:06:18 ; Search time 16.85 seconds
(without alignments)
412.457 Million cell updates/sec

Title: US-09-457-066-2

Perfect score: 1858

Sequence: 1 MSFGLLLTSLAQGRQT.....DVALEHHECDVCVRGTGG 345

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169	9.1	788	1	US-08-572-225-1
2	163.5	8.8	909	3	US-08-936-135-18
3	163.5	8.8	926	3	US-08-936-135-20
4	160.5	8.6	923	3	US-08-936-135-6
5	159.5	8.6	901	3	US-08-936-135-22
6	159.5	8.6	906	3	US-08-936-135-24
7	159.5	8.6	909	3	US-08-936-135-8
8	159.5	8.6	909	3	US-08-936-135-10
9	159.5	8.6	914	3	US-08-936-135-12
10	159.5	8.6	926	3	US-08-936-135-14
11	159.5	8.6	931	3	US-08-936-135-16
12	154	8.3	449	2	US-08-839-008-2
13	154	8.3	449	2	US-08-839-008-9
14	153	8.2	415	4	US-09-032-523-2
15	149	8.0	591	3	US-08-991-408-4
16	149	8.0	1013	2	US-08-866-650-5
17	149	8.0	1013	2	US-09-021-287-5
18	149	8.0	1013	3	US-08-991-408-2
19	148	8.0	1013	2	US-08-866-650-3
20	148	8.0	1013	2	US-09-021-287-3
21	147	7.9	354	4	US-08-915-795-3
22	147	7.6	325	4	US-08-915-795-3
23	140	7.5	321	4	US-08-915-795-3
24	140	7.5	358	4	US-08-915-795-8
25	132	7.1	419	2	US-08-999-811-2
26	132	7.1	419	3	US-09-042-105-2
27	132	7.1	419	3	US-09-042-105-18

Sequence 8, Appli
Sequence 35, Appli
Sequence 5, Appli
Sequence 7, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 11, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 6, Appli
Sequence 6, Appli
Patent No. 5175255
Sequence 4, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 33, Appli

28 132 7.1 419 4 US-08-795-430-8
29 132 7.1 419 4 US-08-510-133A-35
30 128 6.9 401 2 US-08-839-008-5
31 128 6.9 468 2 US-08-839-008-7
32 128 6.9 468 4 US-09-032-523-8
33 126 6.8 419 5 PCT-US96-09001-2
34 125 6.7 415 4 US-08-795-430-11
35 123.5 6.6 277 1 US-08-024-868-2
36 123.5 6.6 277 2 US-08-242-097-2
37 123.5 6.6 277 4 US-09-206-695-2
38 123.5 6.6 277 5 PCT-US96-11995-1
39 121.5 6.5 101 2 US-08-242-097-6
40 121.5 6.5 101 4 US-09-206-695-6
41 116.5 6.3 256 6 5175255-9
42 115 6.2 350 2 US-08-999-811-4
43 115 6.2 350 2 US-08-824-996-2
44 115 6.2 350 3 US-09-042-105-4
45 115 6.2 350 4 US-08-510-133A-33

ALIGNMENTS

RESULT 1
US-08-572-225-1
; Sequence 1, Application US/08572225
; Patent No. 5807981
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojima, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; APPLICANT: Brenner, Mitch
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
; TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/572,225
FILING DATE: 13-DEC-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 8389-031
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-572-225-1

Query Match 9.1%; Score 169; DB 1; Length 788;
Best Local Similarity 39.4%; Pred. No. 1.4e-09;
Matches 43; Conservative 17; Mismatches 39; Indels 10; Gaps 5;

Qy 55 NGSIHSRPHPTYPRTVWRLVA-VENWVIOITFDREGLDEPDDICKYDFVEVEE 113
Db 401 NGSIHSPGWKPEYPPKNCINQWVAPTOYRISLOFDFETEG-----NDVCKYDFVEVRS 455
Qy 114 --PSDGTILGRWCGSGTVPKQISKGNQIRIRFVSDYFVFPSEPGFCIHY 160
Db 456 GLTADSKLHGAFCS-EKPEVITSOYNNMRVEFKSDNTV-SKKGFKAHF 502

RESULT 2

US-08-936-135-18
; Sequence 18, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936.135
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 909 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-18

Query Match 8.8%; Score 163.5; DB 3; Length 909;
Best Local Similarity 32.3%; Pred. No. 6.8e-09;
Matches 43; Conservative 19; Mismatches 64; Indels 7; Gaps 4;

Qy 32 FSSNKEQNGVQDPQHERITITVSTNGSIHSPRPHTYPRTVWRLVAVEENVWIOITFD 91
Db 14 YFSRHQVRGQDPDPCGGRLNSKDGAYITSPGYPDYPSHQNCWIVYAPEPNQKIVLNFN 73
Qy 92 ERFGLEDDEDDICKYDFVEVE--EPSDGTILGRWCGSGTVPKQISKGNQIRIRFVSDY 149
Db 74 PHFEIEKHD---CRYDFIEIRGDSEADLLGKHCGN-IAPPTIISSGSMLYIKFTSD-Y 128
Qy 150 FPSEPGFCIHYNI 162
Db 129 ARQGAGFSLRVEI 141

RESULT 3

US-08-936-135-20
; Sequence 20, Application US/08936135

; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936.135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 926 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-20

Query Match 8.8%; Score 163.5; DB 3; Length 926;
Best Local Similarity 32.3%; Pred. No. 7e-09;
Matches 43; Conservative 19; Mismatches 64; Indels 7; Gaps 4;

Qy 32 FSSNKEQNGVQDPQHERITITVSTNGSIHSPRPHTYPRTVWRLVAVEENVWIOITFD 91
Db 14 YFSRHQVRGQDPDPCGGRLNSKDGAYITSPGYPDYPSHQNCWIVYAPEPNQKIVLNFN 73
Qy 92 ERFGLEDDEDDICKYDFVEVE--EPSDGTILGRWCGSGTVPKQISKGNQIRIRFVSDY 149
Db 74 PHFEIEKHD---CRYDFIEIRGDSEADLLGKHCGN-IAPPTIISSGSMLYIKFTSD-Y 128
Qy 150 FPSEPGFCIHYNI 162
Db 129 ARQGAGFSLRVEI 141

RESULT 4

US-08-936-135-6
; Sequence 6, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA

ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 923 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-6

Query Match 8.6%; Score 160.5; DB 3; Length 923;
Best Local Similarity 36.5%; Pred. No. 1.5e-08;
Matches 42; Conservative 16; Mismatches 50; Indels 7; Gaps 4;
QY 50 INVSTNGSIHSFRPHYTPRNTVLVRLVAEENVVQLTFDERFGLDEDDICKYDFV 109
DB 31 IKNPGYTSCYPHSYHPSEKCEWLIQAPEYQRIINFNPHFDEKRD---CRYDYV 87
QY 110 EV--EPPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDYPPSPGFCIHNYI 162
DB 88 EVIDGENEGRLWKGKCGK-IAPSPVSSGPFIFKIFVSD-YETHGAGFSIRYEI 140

RESULT 5
US-08-936-135-22
Sequence 22, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 901 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-22

Query Match 8.6%; Score 159.5; DB 3; Length 901;
Best Local Similarity 33.1%; Pred. No. 1.8e-08;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;
QY 32 FSSNKEQNGVQDPQHERLIITVSTNGSIHSFRPHYTPRNTVLVRLVAEENVVQLTFD 91
DB 14 YFSGHEVRSQDDPPCGGRPNKSDAGYITSPGYQDPYSHQNCWEIVYAPEPNQKIVLNFN 73
QY 92 ERFGLEDPEDDICKYDFVEVE--EPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDY 149
DB 74 PHEIEKHD---CKYDFEIRDGSEADLLGKHGNG-IAPPTIISSGSVLYIKFTSD-Y 128
QY 150 FPSEPFGFCIHNYI 162
DB 129 ARQGAGFSIRYEI 141

RESULT 6
US-08-936-135-24
Sequence 24, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-24

Query Match 8.6%; Score 159.5; DB 3; Length 906;
Best Local Similarity 33.1%; Pred. No. 1.9e-08;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;

STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UC97-288-2

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 914 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-936-135-12

Query Match 8.6%; Score 159.5; DB 3; Length 914;

Best Local Similarity 33.1%; Pred. No. 1.9e-08;

Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;

QY 32 FSSNKEQNGVQDPQHERIITVSTNGSIHSPRPHTYPRNTVLVRLVAVEENWVQLTFD 91

Db 14 YFSGHEVRSQQDPPCGGRPNKSDAGYITSPGYQDYPHQNCWEIVYAPEPNQKIVLNFN 73

QY 92 ERFGLEDDEDDICKYDFVEVE--EPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDBY 149

Db 74 PHFEIEKHD--CKYDFEIRDGDSEADLLGKHGCGN-IAPPTIISGSLVYIKFTSD-Y 128

QY 150 FPSEPGFCIHNYI 162

Db 129 ARQAGFSLRYEI 141

RESULT 10

US-08-936-135-14

Sequence 14, Application US/08936135

Patent No. 6054293

GENERAL INFORMATION:

APPLICANT: Tessier-Lavigne, Marc

APPLICANT: He, Zhigang

APPLICANT: Chen, Hang

TITLE OF INVENTION: Semaphorin Receptors

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/936,135

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UC97-288-2

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 926 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-936-135-14

Query Match 8.6%; Score 159.5; DB 3; Length 926;

Best Local Similarity 33.1%; Pred. No. 1.9e-08;

Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;

QY 32 FSSNKEQNGVQDPQHERIITVSTNGSIHSPRPHTYPRNTVLVRLVAVEENWVQLTFD 91

Db 14 YFSGHEVRSQQDPPCGGRPNKSDAGYITSPGYQDYPHQNCWEIVYAPEPNQKIVLNFN 73

QY 92 ERFGLEDDEDDICKYDFVEVE--EPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDY 149

Db 74 PHFEIEKHD--CKYDFEIRDGDSEADLLGKHGCGN-IAPPTIISGSLVYIKFTSD-Y 128

QY 150 FPSEPGFCIHNYI 162

Db 129 ARQAGFSLRYEI 141

RESULT 11

US-08-936-135-16

Sequence 16, Application US/08936135

Patent No. 6054293

GENERAL INFORMATION:

APPLICANT: Tessier-Lavigne, Marc

APPLICANT: He, Zhigang

APPLICANT: Chen, Hang

TITLE OF INVENTION: Semaphorin Receptors

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/936,135

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UC97-288-2

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 931 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-16

Query Match 8.6%; Score 159.5; DB 3; Length 931;
Best Local Similarity 33.1%; Pred. No. 1.9e-08;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;
QY 32 FSSNKEQGVQDQHERIITVSTNGSIHPRPHPTYPNTVLVWRLVAEENWVLIQTFD 91
D 14 YFSGEVRSQDDPCGGRPNKDAQYITSPQDYSHQNCWEIVVAPENQKIVLNFN 73
QY 92 ERGLEDDEDDICKYDFVE--EPSDGTILGRWCGSGTVPGKQISGNQIRIRFVSDEY 149
D 74 PHEIEKHDD--CKYDFIEIRDGDSADLIGKHGCGN-IAPPTIISGSLVIKFTSD-Y 128
QY 150 FPSEPFGCIHNYI 162
D 129 AROGAGFSLRIEYI 141

RESULT 12
US-08-839-008-2
; Sequence 2, Application US/08839008
; Patent No. 5916758
; GENERAL INFORMATION:
; APPLICANT: Hurlle, Mark R
; APPLICANT: McDonnell, Peter C
; APPLICANT: McNulty, Dean E
; APPLICANT: Rosen, Craig A
; APPLICANT: Siemens, Ivo R
; APPLICANT: Young, Peter R
; APPLICANT: Yue, Tian-Li
; TITLE OF INVENTION: Smooth Muscle Cell-Derived Migration Factor
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,008
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/563,697
FILING DATE: 28-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50384
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-839-008-2

Query Match 8.3%; Score 154; DB 2; Length 449;

Best Local Similarity 33.8%; Pred. No. 2.6e-08;
Matches 48; Conservative 22; Mismatches 50; Indels 22; Gaps 10;

QY 56 GSIHSPRPHT-YPRNTVLVWRLVAEENWVLIQTFDERFGLDEDDICKYDFVEV--- 111
D 168 GTLTTPNWPESDYPGIGSCSWHIIAPPDQV-IALTFF-EKFDLE--PDTYCRDVSVSFNG 223
QY 112 EEPDGTILGRWCGSGTVPGKQISGNQIRIRFVSDEYPPSPGFCIHNYIWPQT--- 168
D 224 AVSDRRLGLKFGCGD-AVPGSISSEGNELLVQFVSDLSVTAD-GFSASYK-TLPRGTAKE 280
QY 169 -----EAVSPSV-LPDSALP 182
D 281 GCGPGPKRGTEPKVKLPKRSQP 302

RESULT 13
US-08-839-008-9
; Sequence 9, Application US/08839008
; Patent No. 5916758
; GENERAL INFORMATION:
; APPLICANT: Hurlle, Mark R
; APPLICANT: McDonnell, Peter C
; APPLICANT: McNulty, Dean E
; APPLICANT: Rosen, Craig A
; APPLICANT: Siemens, Ivo R
; APPLICANT: Young, Peter R
; APPLICANT: Yue, Tian-Li
; TITLE OF INVENTION: Smooth Muscle Cell-Derived Migration Factor
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,008
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/563,697
FILING DATE: 28-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50384
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-839-008-9

Query Match 8.3%; Score 154; DB 2; Length 449;
Best Local Similarity 33.8%; Pred. No. 2.6e-08;
Matches 48; Conservative 22; Mismatches 50; Indels 22; Gaps 10;

QY 56 GSIHSPRPHT-YPRNTVLVWRLVAEENWVLIQTFDERFGLDEDDICKYDFVEV--- 111
D 168 GTLTTPNWPESDYPGIGSCSWHIIAPPDQV-IALTFF-EKFDLE--PDTYCRDVSVSFNG 223

Search completed: June 29, 2001, 17:09:16
Job time: 178 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:06:48 ; Search time 21.63 Seconds
(without alignments)
1214.989 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 1858
Sequence: 1 MSLLFGLLLTSALAGORQT.....DVALEHHECDVCVRGSTGG 345

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	183.5	9.9	3623	2 T09456	intrinsic factor-B
2	181.5	9.8	730	1 BKH01	procollagen C-endo
3	179.5	9.7	927	1 J00948	A5 antigen precurs
4	176	9.5	707	2 JC2218	procollagen C-endo
5	173	9.3	823	1 A58788	procollagen C-endo
6	171	9.2	3623	2 T08618	intrinsic factor-B
7	169	9.1	986	1 B58788	procollagen C-endo
8	169	9.1	991	2 I49540	procollagen C-endo
9	154	8.3	449	2 A55362	procollagen I C-pr
10	148.5	8.0	1057	1 A39288	dorsal-ventral pat
11	143.5	7.7	686	1 A59271	Ra-reactive factor
12	139.5	7.5	1070	2 T31069	tolluid-BMP-I like
13	138.5	7.5	597	2 S71352	metalloproteinase
14	138	7.4	705	1 C1HURB	complement subcomp
15	137.5	7.4	699	1 I54763	Ra-reactive factor
16	137.5	7.4	1524	2 T30337	polyprotein - Afri
17	133	7.2	1594	2 T30549	hensin - rabbit
18	132	7.1	419	2 S69207	vascular endotheli
19	130.5	7.0	1464	2 S58984	development protei
20	128	6.9	402	2 JH0403	procollagen I C-pr
21	127.5	6.9	767	2 T30018	hypothetical prote
22	127.5	6.9	3871	2 T22812	hypothetical prote
23	125	6.7	198	2 J50735	platelet-derived g
24	123.5	6.6	277	2 A41735	hyaluronate-bindin
25	120.5	6.5	245	1 TVCTSS	platelet-derived g
26	120.5	6.5	275	2 JC6506	tumor necrosis fac
27	118.5	6.4	276	2 A47290	TSG-6 homolog p54
28	114.5	6.2	148	2 D49530	16K vascular endot
29	114.5	6.2	241	1 PFHUG2	platelet-derived g

30	112.5	6.1	200	2 I51551	platelet-derived g
31	112.5	6.1	215	2 S08220	platelet-derived g
32	112.5	6.1	226	2 I51550	platelet-derived g
33	111.5	6.0	319	2 I51569	UWS-2 protein - Af
34	110	5.9	166	2 JN0248	platelet-derived g
35	109	5.9	197	2 S25096	PDGF-related trans
36	108.5	5.8	226	1 TMVYSS	platelet-derived g
37	108	5.8	196	2 A37359	vascular endotheli
38	104.5	5.6	232	2 A41551	platelet-derived g
39	104	5.6	196	2 B28964	platelet-derived g
40	104	5.6	211	1 PFHUG1	platelet-derived g
41	102.5	5.5	133	2 B49530	vascular endotheli
42	102.5	5.5	190	2 S52130	vascular endotheli
43	102.5	5.5	196	2 A48851	platelet-derived g
44	102.5	5.5	1023	2 G96509	protein F27F5_21 l
45	102.5	5.5	2083	2 T42721	CRP-ductin-alpha p

ALIGNMENTS

RESULT 1

T09456

intrinsic factor-B12 receptor Cubilin precursor - human

C:Species: Homo sapiens (man)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T09456

R:Kozayraki, R.; Kristiansen, M.; Silahtaroglu, A.; Hansen, C.; Jacobsen, C.; Tommerup

Blood 91, 3593-3600, 1998

A:Title: The human intrinsic factor-vitamin B12 receptor, cubilin: Molecular character

ion.

A:Reference number: Z16677; MUID:98241400

A:Accession: T09456

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-3623 <KO2>

A:Cross-references: EMBL:AF034611; NID:g3929528; PIDN:AA82612.1; PID:g3929529

C:Genetics:

A:Map position: 10p12

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: receptor; vitamin B12 uptake

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>

F:436-467/Domain: EGF homology <EGF>

Query Match 9.9%; Score 183.5; DB 2; Length 3623;
Best Local Similarity 25.7%; Pred. No. 6.2e-07;
Matches 98; Conservative 46; Mismatches 142; Indels 95; Gaps 24;

QY 10 TSALAGORQ-----GTOAESNLSSKTFQ-----SSNKEQNG-----VODPOHERIT 51

Db 878 SSSILGSPENKKYCGTDIPSFITSVYNFLVTVFVKSSSTENHGFMKFAEDLACGEILT 937

QY 52 VSTNGSIHSRPFHTYPRNTVLVRLVAVENWVLIQTDFERFCELEDPEDDICKYDFEV 111

Db 938 EST-GTIQSPGHVNVPHGINCTWHIL-VQPNHLIHLMF-ETFHLEFHYN--CTNDYLEV 992

QY 112 EPSPDGTILGRWCGSGTVPGKQISKNQIRIRFVSODEYFPSEPGFCIHYNIV-----MP 165

Db 993 YDTSSETSLGRYCGK-SIPPSLTSSGNSLMVFTVDSDLAYE-GFLINYAISATACLQ 1050

QY 166 QTEAVSPSVLPSPALPLDLNNNAITAFSTLEDLIRYLEPERMOLDLEDLYR----PTWQL 222

Db 1051 DVTDDLGTFTSP-----NFPNN-----YNNW---ECIVRITVRGTOL 1085

QY 223 LGKAFVGRKSRVVDNLNLTLE--EVR-----LY--SCTPRNFVSIREELK-R 265

Db 1086 IAVHFTNFSLEEAIG-NYYTDFLEIRDGGYKESPLGIFGYSNLPLPTIISHSNKLWLKFK 1144

QY 266 TDTI-----FWPGCLLVKRGCGNCCACLNHCNEQCQVPSKVTKYHE---VLQLRPK 314

Db 1145 SDQIDTRSGFSAYWDGS--STGCGGN----LTTSSGTFITSPNPMYPYHSSECYWLKSS 1198

QY 315 TGVRLHLSLTDVALEHHEEC 335
DB 1199 HG-SAFEFEKDFLHHPNC 1218

RESULT 2
BMH01
procollagen C-endopeptidase (EC 3.4.24.19) precursor, splice form BMP1 - human
N:Alternate names: bone morphogenic protein 1 (BMP1)
C:Species: Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
C:Accession: A37278; E58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hew
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730
A:Accession: A37278
A:Molecule type: mRNA
A:Residues: 1-730 <WOZ>
A:Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500
C:Genetics:
A:Gene: GDB:BMP1
A:Cross-references: GDB:125203; OMIM:112264
A:Map position: 8p21-8p21
C:Function:
A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology; F
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; gl
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-730/Product: procollagen C-endopeptidase splice form BMP1 #status predicted <MAT>
F:130-321/Domain: astacin homology <AST>
F:322-431/Domain: C1r/C1s repeat homology <C1R1>
F:435-544/Domain: C1r/C1s repeat homology <C1R2>
F:551-587/Domain: EGF homology <EGF>
F:591-700/Domain: C1r/C1s repeat homology <C1R3>
F:91,142,332,363,599/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:163-319,185-205,348,375-397,435-461,488-510,551-559,572,574-587,591-617,644-66
F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.8%; Score 181.5; DB 1; Length 730;
Best Local Similarity 37.3%; Pred. NO. 1.2e-07;
Matches 50; Conservative 20; Mismatches 47; Indels 17; Gaps 7;

QY 55 NGSISHPRPHTYPRNTVLVRLVA-VEENVWQLTFDERFGLDPEDDICKYDFVEVEE 113
DB 599 NGSITSPGWPKEYPPNKNKCIWQVAPTQYRISLQDFEFTEG-----NDVCKYDFVEVRS 653

QY 114 --PSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDYFPSEPGFCIHYNLVMPQFTEAV 171
DB 654 GLTADSKLHGKFCGS-EKPEVITSQIAHMRVEFKSDNTV-SKKGFRHF-----FSEK- 704

QY 172 SPSVLPPSPALPLDL 185
DB 705 RPALQPPRGRPHQL 718

RESULT 3
JQ0948
A5 antigen precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JH0466; JQ0948
R:Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A:Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homology
A:Reference number: JH0466; MUID:91337458
A:Accession: JH0466
A:Molecule type: mRNA
A:Residues: 1-927 <TAK>

A:Cross-references: GB:D10467; GB:D01077; NID:g222962; PIDN:BA01260.1; PID:g222963
A:Experimental source: tadpole, brain
A:Note: This protein has motifs homologous to complement components C1r and C1s and t
C:Comment: This protein is a neuronal cell surface molecule involved in the neuronal
C:Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-termina
C:Keywords: duplication; glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-927/Product: A5 antigen #status predicted <ASA>
F:27-138/Domain: C1r/C1s repeat homology <C1R1>
F:147-262/Domain: C1r/C1s repeat homology <C1R2>
F:274-424/Domain: discoidin I amino-terminal homology <DN1>
F:430-584/Domain: discoidin I amino-terminal homology <DN2>
F:646-812/Domain: MAM homology <MAM>
F:861-883/Domain: transmembrane #status predicted <TMM>
F:150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 9.7%; Score 179.5; DB 1; Length 927;
Best Local Similarity 31.6%; Pred. NO. 2.4e-07;
Matches 55; Conservative 29; Mismatches 73; Indels 17; Gaps 7;

QY 50 ITVSTNGSIHSPRPHTYPRNTVLVRLVAVEENVWQLTFDERFGLDPEDDICKYDFV 109
DB 31 IKITSPSVLTSAGYPHSYPPSORCEWLQAPHYQIRIMFNPNPHELDRE---CKYDIV 87

QY 110 EV--EEPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDYFPSEPGFCIHYNVMP-- 165
DB 88 EVIDGDNANGQLLKCYCGK-IAPSLVSTGPSIFIRFVSDYETPG-AGFSIRYEFKTPG 145

QY 166 ---QFTEA--VSPSLPPSPALPLDLNNAITAFSTLEDLIRLPEPERWOLDLE 213
DB 146 ECSRNTSSNGVIKSPKPYKPYNALECTYIIFAPKMQEIV--LEFSEFELEAD 197

RESULT 4
JC2218
procollagen C-endopeptidase (EC 3.4.24.19) - African clawed frog (clone 22AN)
N:Alternate names: bone morphogenic protein 1
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Sep-1999
C:Accession: JC2218
R:Maeno, M.; Xue, Y.; Wood, T.I.; Ong, R.C.; Kung, H.
Gene 134, 257-261, 1993
A:Title: Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic pr
A:Reference number: JC2218; MUID:94085787
A:Accession: JC2218
A:Molecule type: mRNA
A:Residues: 1-707 <MAE>
A:Cross-references: GB:L12249; NID:g406540; PIDN:AAA16313.1; PID:g406541
C:Comment: This protein induces ectopic cartilage formation in vivo.
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology
C:Keywords: beta-hydroxyasparagine; glycoprotein; hydrolase; metalloproteinase; zinc
F:93-284/Domain: astacin homology <AST>
F:285-397/Region: complement 1r/1s-like repeat
F:285-394/Domain: C1r/C1s repeat homology <C1R1>
F:398-510/Region: complement 1r/1s-like repeat
F:398-507/Domain: C1r/C1s repeat homology <C1R2>
F:514-550/Domain: EGF homology <EGF>
F:554-666/Region: complement 1r/1s-like repeat
F:554-663/Domain: C1r/C1s repeat homology <C1R3>
F:62,105,295,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:176,180,186,235/Binding site: zinc (His, His, His, Tyr) #status predicted
F:177/Active site: Glu #status predicted
F:528/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.5%; Score 176; DB 2; Length 707;
Best Local Similarity 43.3%; Pred. NO. 3.5e-07;
Matches 45; Conservative 15; Mismatches 36; Indels 8; Gaps 5;

QY 55 NGSISHPRPHTYPRNTVLVRLVAVEENVWQLTFDERFGLDPEDDICKYDFVEVEE- 113
DB 562 NGSINSPGWPKEYPPNKNKCIWQVAPTQ-YRISLKFQD---FETEGNDVCKYDFVEVRS 617

```
QY 114 -PSDITILGRWCSGTVPCKQISKGNOIRIRVSDVEYFPSEPGF 156
      ||| : : ||| : ||| : ||| : ||| : ||| : |||
Db 618 LTSDKLLHGKFCGS-ELPAVITSOYNNMRIEFKSDNTV-SKKGF 659
```

RESULT 5

A58788
collagen C-endopeptidase (EC.3.4.24.19) precursor, splice form HIS - human
N;Alternate names: bone morphogenic protein splice form BMP-1/HIS
C;Species: Homo sapiens (man)
C;Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 18-Jun-1999
C;Accession: A37278; A58788
R;Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz
Science 242, 1528-1534, 1998
A;Title: Novel regulators of bone formation: molecular clones and activities.
A;Reference number: A37278; PMID:89072730

A;Accession: A3728
A;Molecule type: mRNA
A;Residues: 1-702, 'EKRPALQPPRRGHQLKFRVQKNRTTPQ' <NO>
A;Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500
R;Takahara, K.; Lyons, G.E.; Greenspan, D.S.
J. Biol. Chem. 269, 32572-32578, 1994
A;Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue
A;Reference number: A58788; MUID:95096114
A;Accession: A58788
A;Molecule type: mRNA
A;Residues: 703-823 <TAK>
A;Cross-references: GB:L35278; NID:g619423; PIDN:AAC41703.1; PID:g619424
C;Genetics:
A;Gene: GDB:BMPI, BMP-1
A;Cross-references: GDB:L35203; OMIM:112264
A;Map position: 8p21-8p21

C:Function:
A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type I
C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology; F
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; gl
F:1-22/Domain: signal sequence #status predicted <SIG>
F:2-823/Product: procollagen C-endopeptidase splice form HIS #status predicted <MAT>
F:130-331/Domain: astacin homology <AST>
F:322-431/Domain: C1r/C1s repeat homology <CLR1>
F:435-544/Domain: C1r/C1s repeat homology <CLR2>
F:551-587/Domain: EGF homology <EGF>
F:591-700/Domain: C1r/C1s repeat homology <CLR3>
F:738-752/Region: histidine-rich
F:91, 142, 332, 363, 599/Binding site: carbohydrate (Asn) #status predicted
F:163-319, 385-205, 322-348, 375-397, 435-461, 488-510, 551-563, 559-572, 574-587, 591-617, 644-66
F:213-217, 223, 272/Binding site: zinc (His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.3%; Score 173; DB 1; Length 823;
Best Local Similarity 35.6%; Pred. No. 7.7e-07;
Matches 52; Conservative 19; Mismatches 49; Indels

[illegible]

RESULT

T08618
intrinsic factor-B12 receptor CUBILIN precursor - rat
C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08618
R:Moestrup, S.K.; Kozyraki, R.; Kristiansen, M.; Kayser, J.H.; Rasmussen, F.
J. Biol. Chem. 273, 5235-5242, 1998
A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic
A:Reference number: Z16459; MUID:98148073

A: Status: preliminary: translated from GB/EMBL/pdb

A; Molecule type: mRNA

A:Residues: 1-3623 <MOE>
A:Cross-references: EMBL:AF022247; NID:g3834379; PIDN:AAC71661.1; PID:g3834380
C:Genetics:

A; Gene: CUBILIN

C; Superfamily: unassigned EGF-related proteins; EGF homology

C;Keywords: egg yolk; endocytosis; glycoprotein; intest

F.1-20/Domain: signal sequence status predicted <SIG>
F.21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
F.133-164/Domain: EGF homology <EGF>
F.436-467/Domain: EGF homology <EGF>

Query Match 9.2%: score 171: DB 2: length 3623:

Best Local Similarity 24.98; Pred. No. 7.5e-06;

Matches	92;	Conservative	52;	Mismatches	138;	Indels	88;	Gaps	23;
---------	-----	--------------	-----	------------	------	--------	-----	------	-----

QY 2 SLFGLLLTSALAGOROGTOAESNLSKFOFSSNKEONGVODPOHERIITVSTNGSIHSP 61

Db 901 SVYNILYVTFVKSSSME-----NRGTTAKFSSDKLECG-----EVLTASt-GIESP 946

QY 62 RPH TYP RNTVLVWRLVAVEENVW IQLTFDERFGL EDPEDDICKYDFVEVEEPSDGTILG 121

Db 947 GHPNVYPRGVNCTWHVV-VQRGQLIRLEFSS-FYLEFHYN--CTNDYLEIYDTAAQTFGL 1002

QV 122 RWCGSGTVPGKOISKGNQIRIRFVSDEYFPSEPGFCIHY-----NIVMPOFTEAVSPSV 175

Db 1003 RYCGK-SIPPTSNSNSKLIYVSDSALAHE-GFSINYEADASSVCLDYDTDFEGMLS 1060

QY 176 LP--PSALPLDL-----LNNAI---TAFSTLED---LIRYLEPERWQDLLEDL 215

[illegible]

Qv	216	YRPTWOLIGKAEFVEGRKSRVVDINLTVEEVLY-----SCTPRNFESVSIRPEIKRPTDT	268
DB	1061	SPNEFPNNYPSNWECIYRITVGLNQIALHF.TDF--TLEDYFGSQCVDFE-----IRDG	1111

Db 1113 GYETSPVVG---IY-CGSVLPPTTISHSNKWLKFKSDAALTAKGFSA-----1156

QY 269 IFWP GCLLVKRCGGNCACCLHNCNEQCVP SKVTKKYE---VLQLRPKTGVRGLHKS LT 325

QW 326 DVATLEUUEEC 325
Db 1157 -YWDGS--STGCGGN----LTTPTGVLTSPPNPMYPYHSSECYWRLEASHG-SPFELEFQ 1208

QY	326	DVALEHHEC	333

Db 1209 DFHLEHHPSC 1218

RESULT

procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - human
B58788
N:Alternate names: bone morphogenic protein 1, tolloid-like splice form
C:Species: Homo sapiens (man)
C:Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 16-Jul-1999
C:Accession: A37278; B58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsuoka, L.M.; Whitters, M.J.; Kriz, R.W.;
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730

A; Accession: A37278

A;Molecule type: mRNA

A;Residues: 1-702, 'EKRPALQPPRGRPHQLKFRVQKRNPQ' <WOZ>

A;Cross-references: GB:M22488; NID:g179499;

R; Takahara, K.; Lyons, G.E.; Greenspan	
I Biol Chem	269 32572-32578 1994

J. BIOL. CHEM. 269, 32372-32376, 1994

A:Reference number: A58788: MUID:95096114

A:Accession: B58788
A:Molecule type: mRNA
A:Residues: 703-986 <TAK>
A:Cross-references: GB:L35279; NID:g619860; PIDN:AA41710.1; PID:g619861
C:Genetics:
A:Gene: GDB:BMP1; BMP-1
A:Cross-references: GDB:125203; OMIM:112264
A:Map position: 8p21-8p21
C:Function:
A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type I
C:Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology; F:
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; g1
F:122/Domain: signal sequence #status predicted <SIG>
F:23-986/Product: procollagen C-endopeptidase tolloid-like splice form #status predicted
F:130-321/Domain: astacin homology <AST>
F:322-431/Domain: Clr/Cls repeat homology <CLR1>
F:435-544/Domain: Clr/Cls repeat homology <CLR2>
F:551-587/Domain: EGF homology <EG1>
F:591-700/Domain: Clr/Cls repeat homology <CLR3>
F:707-742/Domain: EGF homology <EG2>
F:747-856/Domain: Clr/Cls repeat homology <CLR4>
F:860-973/Domain: Clr/Cls repeat homology <CLR5>
F:91-142-332-363-599/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:163-319-185-205-322-348-375-397-435-461-488-510-551-563-559-572-574-587-591-617-644-66
F:213-217-223-272/Binding site: zinc (His, His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565-720/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.1%; Score 169; DB 1; Length 986;
Best Local Similarity 39.4%; Pred. No. 2.1e-06;
Matches 43; Conservative 17; Mismatches 39; Indels 10; Gaps 5;
QY 55 NGSIHSPRPHTYPRNTVLVWRLVA-VEENVWLTQTFDERFGLDEPDICKYDFVEVEE 113
DB 599 NGSITPGWPKETPPNKNKIWLQVAPTOYRISLQDFEFTEG-----NDVKYDFVEVRS 653
QY 114 --PSDGTILGRWCGSGTVPKGKQISGNQIRIRFVSDEYFPSPGFCIH 160
DB 654 GLTADSKLHGKFCGS-EKPEVITSQYNNMRVEFKSDNTV-SKKGFKAHF 700

RESULT 8
I49540
procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: I49540.
R:Fukagawa, M.; Noboru, S.; Hogan, B.L.M.; Jones, C.M.
Dev. Biol. 163, 175-183, 1994
A:Title: Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1) which is rel
A:Reference number: I49540; MUID:94229342
A:Accession: I49540
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-991 <RES>
A:Cross-references: GB:L24755; NID:g439606; PIDN:AAA37306.1; PID:g439607
C:Genetics:
A:Gene: Bmp-1
C:Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology; B
C:Keywords: hydrolase; metalloproteinase; zinc
F:135-326/Domain: astacin homology <AST>
F:556-592/Domain: EGF homology <EG1>
F:596-705/Domain: Clr/Cls repeat homology <CLR>
F:712-747/Domain: EGF homology <EG2>
F:218-222-228-277/Binding site: zinc (His, His, His, Tyr) #status predicted
F:219/Active site: Glu #status predicted

Query Match 9.1%; Score 169; DB 2; Length 991;
Best Local Similarity 39.4%; Pred. No. 2.2e-06;
Matches 43; Conservative 17; Mismatches 39; Indels 10; Gaps 5;

QY 55 NGSIHSPRPHTYPRNTVLVWRLVA-VEENVWLTQTFDERFGLDEPDICKYDFVEVEE 113
DB 604 NGSITPGWPKETPPNKNKIWLQVAPTOYRISLQDFEFTEG-----NDVKYDFVEVRS 658
QY 114 --PSDGTILGRWCGSGTVPKGKQISGNQIRIRFVSDEYFPSPGFCIH 160
DB 659 GLTADSKLHGKFCGS-EKPEVITSQYNNMRVEFKSDNTV-SKKGFKAHF 705

RESULT 9
A55362
procollagen I C-proteinase enhancer protein precursor - human
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: A55362
R:Takahara, K.; Kessler, E.; Biniaminov, L.; Brusel, M.; Eddy, R.L.; Jani-Sait, S.; S
J. Biol. Chem. 269, 26280-26285, 1994
A:Title: Type I procollagen COOH-terminal proteinase enhancer protein: identification
A:Reference number: A55362; MUID:95014462
A:Accession: A55362
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-449 <TAK>
A:Cross-references: GB:L33799; NID:g642907; PIDN:AAA61949.1; PID:g642908
C:Genetics:
A:Gene: GDB:PCOLCE
A:Cross-references: GDB:305468; OMIM:600270
A:Map position: 7q21.3-7q22
C:Superfamily: Clr/Cls repeat homology
C:Keywords: extracellular protein; glycoprotein; pyroglutamic acid
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-449/Product: #status predicted <MAT>
F:37-146/Domain: Clr/Cls repeat homology <CLR1>
F:159-270/Domain: Clr/Cls repeat homology <CLR2>
F:26/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:29-431/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.3%; Score 154; DB 2; Length 449;
Best Local Similarity 33.8%; Pred. No. 1.6e-05;
Matches 48; Conservative 22; Mismatches 50; Indels 22; Gaps 10;
QY 56 GSIHSPRPHTYPRNTVLVWRLVA-VEENVWLTQTFDERFGLDEPDICKYDFVEV--- 111
DB 168 GLTTPNPPESDYPGICSSWHIAPPDQV-TALTF-EKFDLE--PDTCRYDSVSFVNG 223
QY 112 EPSPDGTILGRWCGSGTVPKGKQISGNQIRIRFVSDEYFPSPGFCIHNIWMPQFT--- 168
DB 224 AVSDSRRRLGKFCGD-AVPGSISSEGNELLQVFDLSVTAD-GFSASYK-TLPRGTAK 280
QY 169 -----EAVSPSV-LPPSALP 182
DB 281 GGGPGPKRGCTEPKVKLPPKSQLP 302

RESULT 10
A39288
dorsal-ventral patterning protein tolloid (EC 3.4.24.-) - fruit fly (Drosophila melan
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39288
R:Shimell, M.J.; Ferquson, E.L.; Childs, S.R.; O'Connor, M.B.
Cell 67, 469-481, 1991
A:Title: The Drosophila dorsal-ventral patterning gene tolloid is related to human bo
A:Reference number: A39288; MUID:92034970
A:Accession: A39288
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1057 <SHI>
A:Cross-references: GB:M76976; NID:g157305; PIDN:AAA28491.1; PID:g157306
C:Genetics:
A:Gene: FlyBase:tlid
A:Cross-references: FlyBase:FBgn0003719

A;Residues: 1-705 <LEV>
A;Accession: A24170
A;Molecule type: mRNA
A;Cross-references: GB:M14058; NID:gl79643; PIDN:AA51851.1; PID:gl79644
R;Journet, A.; Tosi, M.
Biochem. J. 240, 783-787, 1986
A;Title: Cloning and sequencing of full-length cDNA encoding the precursor of human comp
A;Reference number: A29768; MUID:87156625
A;Accession: A29768
A;Molecule type: mRNA
A;Residues: 1-151, 'L', 153-705 <JOU>
A;Cross-references: GB:X04701; NID:g29538; PIDN:CAA28407.1; PID:g29539
R;Arlaud, G.J.; Willis, A.C.; Gagnon, J.

3

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:08:48 ; Search time 12.01 Seconds
(without alignments)
984.025 Million cell updates/sec

Title: US-09-457-066-2

Perfect score: 1858

Sequence: 1 MSLLFGLLLTSALAQORQT.....DVALEHHECDVCVRGSTGG 345

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	179.5	9.7	928	1 NRPI_XENLA	P28824 xenopus lae
2	176	9.5	707	1 BMP1_XENLA	P98070 xenopus lae
3	173.5	9.3	616	1 SPAN_STRPU	P98068 strongyloce
4	169	9.1	986	1 BMP1_HUMAN	P13497 homo sapien
5	169	9.1	991	1 BMP1_MOUSE	P98063 mus musculu
6	163.5	8.8	922	1 NRPI_RAT	Q9qwj9 rattus norv
7	163.5	8.8	931	1 NRP2_HUMAN	O60462 homo sapien
8	163	8.8	925	1 NRP2_RAT	O35276 rattus norv
9	160.5	8.6	923	1 NRP1_MOUSE	P97333 mus musculu
10	159.5	8.6	931	1 NRP2_MOUSE	O35375 mus musculu
11	158.5	8.5	923	1 NRP1_HUMAN	O44786 homo sapien
12	157	8.4	914	1 NRPI_CHICK	P79795 gallus gall
13	154	8.3	449	1 PCOL_HUMAN	Q51113 homo sapien
14	148.5	8.0	1057	1 TLD_DROME	P25723 drosophila
15	143.5	7.7	686	1 MAS2_HUMAN	O00187 homo sapien
16	139.5	7.5	597	1 BP10_PABLI	P42674 paracentrot
17	138.5	7.5	704	1 CRAR_MOUSE	P98064 mus musculu
18	138	7.4	639	1 BMP1_STRPU	P98069 strongyloce
19	138	7.4	705	1 C1R_HUMAN	P00736 homo sapien
20	137.5	7.4	699	1 CRAR_HUMAN	P48740 h complemen
21	132	7.1	419	1 VEGC_HUMAN	P49767 homo sapien
22	128	6.9	468	1 PCOL_MOUSE	Q61398 mus musculu
23	128	6.9	468	1 PCOL_RAT	O08628 rattus norv
24	125	6.7	213	1 PDGA_RABIT	P34007 oryctolagus
25	125	6.7	415	1 VEGC_MOUSE	P97953 mus musculu
26	123.5	6.6	277	1 TSG6_HUMAN	P98066 homo sapien
27	120.5	6.5	245	1 PDGB_FELCA	P12919 felis silve
28	118.5	6.4	276	1 TSG6_RABIT	P98065 oryctolagus
29	114.5	6.2	148	1 VEGH_ORFN7	P52585 orf virus (
30	114.5	6.2	241	1 PDGB_HUMAN	P01127 homo sapien
31	112.5	6.1	226	1 PDGA_XENLA	P13698 xenopus lae
32	111.5	6.0	514	1 UVS2_XENLA	P42664 xenopus lae
33	110	5.9	211	1 PDGA_MOUSE	P20033 mus musculu

34	109	5.9	204	1 PDGA_RAT	P28576 rattus norv
35	108.5	5.8	164	1 VEGF_CAVPO	P26617 cavia porce
36	108.5	5.8	226	1 TSIS_SMSAV	P01128 simian sarc
37	105.5	5.7	188	1 VEGB_HUMAN	P49765 homo sapien
38	104.5	5.6	215	1 VEGF_HUMAN	P15692 homo sapien
39	104	5.6	211	1 PDGA_HUMAN	P04085 homo sapien
40	102.5	5.5	133	1 VEGH_ORFN2	P52584 orf virus (
41	102.5	5.5	190	1 VEGF_PIG	P49151 sus scrofa
42	101.5	5.5	146	1 VEGF_SHEEP	P50412 ovis aries
43	101.5	5.5	190	1 VEGF_BOVIN	P15691 bos taurus
44	101.5	5.5	225	1 PDGB_RAT	Q05028 rattus norv
45	98.5	5.3	241	1 PDGB_SHEEP	Q95229 ovis aries

ALIGNMENTS

RESULT 1
ID NRPI_XENLA STANDARD; PRT; 928 AA.
AC P28824:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN) (A5 ANTIGEN).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91337458; PubMed=1908252;
RA Takagi S., Hirata T., Agata K., Mochil M., Eguchi G., Fujisawa H.;
RT "The A5 antigen, a candidate for the neuronal recognition molecule,
has homologues to complement components and coagulation factors.";
RL Neuron 7:295-307(1991).
CC -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
SYSTEM. IT MEDIATES THE CHEMOPULSANT ACTIVITY OF SEMAPHORINS (BY
SIMILARITY). PRESUMED TO BE INVOLVED IN THE NEURONAL RECOGNITION
BETWEEN THE OPTIC NERVE FIBERS AND THE VISUAL CENTERS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: RETINAL GANGLION CELLS AND VISUAL CENTER
NEURONS.
CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10467; BAA01260.1; -
CC InterPro; IPR000421; -
CC InterPro; IPR000859; -
CC InterPro; IPR000998; -
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00629; MAM; 1.
CC PRINTS; PR00020; MAMDOMAIN.
CC PROSITE; PS00740; MAM_1; 1.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS50060; MAM_2; 1.

```
KW Transmembrane; Glycoprotein; Neurexin; Signal; Repeat; Receptor;
FT Antigen.
FT SIGNAL 1 21
FT CHAIN 22 928
FT DOMAIN 22 860
FT TRANSMEM 861 883
FT DOMAIN 884 928
FT DOMAIN 27 141
FT DOMAIN 147 265
FT DOMAIN 275 424
FT DOMAIN 431 584
FT DOMAIN 646 812
FT DISULFID 27 54
FT DISULFID 82 104
FT DISULFID 147 173
FT DISULFID 206 228
FT DISULFID 275 424
FT DISULFID 431 584
FT CARBOHYD 150 150
FT CARBOHYD 261 261
FT CARBOHYD 300 300
FT CARBOHYD 523 523
FT CARBOHYD 844 844
SQ SEQUENCE 928 AA; 103416 MW; AF6B323B0A4C789D CRC64;

Query Match 9.7%; Score 179.5; DB 1; Length 928;
Best Local Similarity 31.6%; Pred. No. 1.4e-07;
Matches 55; Conservative 29; Mismatches 73; Indels 17; Gaps 7;

QY 50 ITVTNGSIHSPRPHYPRTVTLVRLVAVENWVQLTDFERFGLDEPDCKYDFV 109
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
D 31 IKITSPSYLTAGYSPHSPYPPSQRCEWLIQAPEHYQIRIMENPHFDLEDE---CKYDV 87
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

QY 110 EV--EESDGTILGRWCGSTVPGKQISKGNIQIRFVSDYFSEPGFCIHYNVMP-- 165
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
D 88 EVIDGDNAGOLLGKYCKGK-IAPSLVSTGPSIFIRFVSDYETPG--LEFSEFELEAD 145
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

QY 166 ----QFTEA--VSPSVLPSPALPDLLNNAITAFSTLEDLRLYPEPERWQDLE 213
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
D 146 ECSNFSSNGVIRKPYKPYNALECTYIIFAPKMQEIV--LEFSEFELEAD 197
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

RESULT 2
BMP1_XENLA STANDARD; PRT; 707 AA.
ID BMP1_XENLA STANDARD; PRT; 707 AA.
AC P98070;
DT 01-FEB-1996 (Rel. 33, Created)
DE 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.-) (BMP-1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94085787; PubMed=8262384;
RA Maeno M., Xue Y., Wood T.I., Ong R.C., Kung H.F.;
RT "Cloning and expression of cDNA encoding Xenopus laevis bone
RT morphogenetic protein-1 during early embryonic development.";
RL Gene 134:257-261(1993).
CC -!- FUNCTION: INVOLVED IN PATTERN FORMATION IN GASTRULA AND LATER
CC DIFFERENTIATION OF DEVELOPING ORGANS.
CC -!- DEVELOPMENTAL STAGE: BLASTULA, EARLY GASTRULA AND HATCHED
CC TADPOLES; LITTLE OR NO EXPRESSION IN MORULA AND LATE GASTRULA.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 CUB DOMAINS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HMBL; L12249; AAL16313.1;
DR HSSP; P00736; LAPO.
DR MEROPS; M12.005;
DR InterPro; IPR000130;
DR InterPro; IPR000152;
DR InterPro; IPR000561;
DR InterPro; IPR000859;
DR InterPro; IPR001506;
DR InterPro; IPR001881;
DR Pfam; PF01400; Astacin; 1.
DR Pfam; PF00431; CUB; 3.
DR Pfam; PF00008; EGF; 1.
DR PRINTS; PRO0480; ASTACIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS01180; CUB; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase; Protease;
KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
KW Glycoprotein.
FT SIGNAL 1 ?
FT PROPEP ? 83
FT CHAIN 84 707
FT DOMAIN 84 284
FT DOMAIN 285 397
FT DOMAIN 398 509
FT DOMAIN 510 551
FT DOMAIN 554 666
FT METAL 176 176
FT ACT_SITE 177 177
FT METAL 180 180
FT METAL 186 186
FT DISULFID 514 526
FT DISULFID 522 535
FT DISULFID 537 550
FT CARBOHYD 62 62
FT CARBOHYD 105 105
FT CARBOHYD 295 295
FT CARBOHYD 326 326
FT CARBOHYD 562 562
SQ SEQUENCE 707 AA; 80673 MW; 1B6980D716DC9B8D CRC64;

Query Match 9.5%; Score 176; DB 1; Length 707;
Best Local Similarity 43.3%; Pred. No. 2e-07;
Matches 45; Conservative 15; Mismatches 36; Indels 8; Gaps 5;

QY 55 NGSIHSPRPHYPRTVTLVRLVAVENWVQLTDFERFGLDEPDCKYDFVEVEE- 113
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
D 562 NGSINSGPWPKYPPNKNICWLQVAPTO-YRISLKFQD---FETEGNDVCKYDFVEVRSG 617
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

QY 114 -PSDGTILGRWCGSTVPGKQISKGNIQIRFVSDYFSEPGF 156
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
D 618 LTSDSKLHGKFCGS-ELPAVITTSQYNNMRFKSDNTV-SKKG 659

RESULT 3
ID SPAN_STRPU STANDARD; PRT; 616 AA.
AC P98068;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
CC -----
```



```
FT SIGNAL 1 21 POTENTIAL..
FT CHAIN 22 922 NEUROFILIN-1.
FT DOMAIN 22 855 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 856 920 POTENTIAL.
FT DOMAIN 881 922 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 841 841 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 922 AA; 103082 MW; CC6F82AD098B0F2E CRC64;

Query Match 8.8%; Score 163.5; DB 1; Length 922;
Best Local Similarity 36.5%; Pred. No. 3.4e-06;
Matches 42; Conservative 16; Mismatches 50; Indels 7; Gaps 4;

QY 50 INVSTNGSHSPRFTPTNTVLVRLVAENVVLIQTFDERFGLDEDDICKYDFV 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 31 IKIENPGYLTSPGYSHYSEKCEWLQIAPEYQIRIMINFNPHFDLDRD---CKYDYV 87
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 110 EV--EPPSGDTILGRWCGTVPKGKQISGNQIRIRFVSDEYFPPSEPGCIHNYI 162
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 88 EVIDGENEGRWNGKFCGR-IAPSPVSSGPFLEIKFVSD-YETHGAGFSIRYEI 140
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
NRP2_HUMAN
ID NRP2_HUMAN STANDARD; PRT; 931 AA.
AC O60462; O14820; O14821;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE NEUROFILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165
DE RECEPTOR 2).
GN NRP2 OR VEGF165R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A0 AND A17).
RX MEDLINE=97470888; PubMed=9331348;
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
RT "Neurofilin-2, a novel member of the neurofilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III.";
RL Neuron 19:547-559(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A22).
RX TISSUE=Breast;
RC MEDLINE=98188099; PubMed=9529250;
RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neurofilin-1 is expressed by endothelial and tumor cells as an
RT isoform-specific receptor for vascular endothelial growth factor.";
RL Cell 92:735-745(1998).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20309748; PubMed=10748121;
RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
RT "Neurofilin-2 and neurofilin-1 are receptors for the 165-amino acid
RT form of vascular endothelial growth factor (VEGF) and of placenta
```

```
RT growth factor-2, but only neurofilin-2 functions as a receptor for
RT the 145-amino acid form of VEGF.";
RL J. Biol. Chem. 275:18040-18045(2000).
CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165
CC AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF. WITH
CC -1- SUBUNIT: NEUROFILIN-2 PROBABLY FORMS A HETEROMERIC COMPLEX WITH
CC NEUROFILIN-1 IN ORDER TO BE A FUNCTIONAL SEMAPHORIN 3C RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A22 (SHOWN HERE), A0 AND A17;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022859; AAC51788.1; -
DR EMBL; AF022860; AAC51789.1; -
DR EMBL; AF016098; AAC12922.1; -
DR MIM; 602070; -
DR InterPro; IPR000421; -
DR InterPro; IPR000859; -
DR InterPro; IPR000998; -
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS50060; MAM_2; 1.
DR Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL 1 20 OR 22 (POTENTIAL).
FT CHAIN 21 931 NEUROFILIN-2.
FT DOMAIN 21 864 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 865 889 POTENTIAL.
FT DOMAIN 890 931 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 142 CUB 1.
FT DOMAIN 149 267 CUB 2.
FT DOMAIN 277 427 F5/8 TYPE C 1.
FT DOMAIN 434 592 F5/8 TYPE C 2.
FT DOMAIN 642 802 MAM.
FT DOMAIN 671 674 POLY-SER.
FT DISULFID 28 55 BY SIMILARITY.
FT DISULFID 83 105 BY SIMILARITY.
FT DISULFID 149 175 BY SIMILARITY.
FT DISULFID 208 230 BY SIMILARITY.
FT DISULFID 277 427 BY SIMILARITY.
FT DISULFID 434 592 BY SIMILARITY.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 809 813 MISSING (IN ISOFORM A17).
FT VARSPLIC 809 830 MISSING (IN ISOFORM A0).
FT CONFLICT 602 602 E -> K (IN REF. 1).
SQ SEQUENCE 931 AA; 104830 MW; 270CBAB69A0A797C CRC64;

Query Match 8.8%; Score 163.5; DB 1; Length 931;
Best Local Similarity 32.3%; Pred. No. 3.4e-06;
Matches 43; Conservative 19; Mismatches 64; Indels 7; Gaps 4;

QY 32 FSSNKGNGVQDPOHERIIITVSTNGSHSPRFTPTNTVLVRLVAENVVLIQTFD 91
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 14 YFSRHQVRGQDPDPCGGRLNSKDAGYITSPGPDQYPSHQNCWEIVYAPEPNQIVLN 73
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

FT	DISULFID	28	55	BY SIMILARITY.
FT	DISULFID	83	105	BY SIMILARITY.
FT	DISULFID	149	175	BY SIMILARITY.
FT	DISULFID	208	230	BY SIMILARITY.
FT	DISULFID	277	427	BY SIMILARITY.
FT	DISULFID	434	592	BY SIMILARITY.
FT	CARBOHYD	152	152	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	157	157	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	629	629	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	833	833	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	834	834	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	925 AA; 103896 MW; 3BF62903F644851C CRC64;		

Query Match 8.8%; Score 163; DB 1; Length 925;
 Best Local Similarity 33.8%; Pred. No. 3.7e-06;
 Matches 45; Conservative 20; Mismatches 60; Indels 8; Gaps

QY	32	FSSNKEONGVQDPOHERLIITVSTNGSIHSRPFHTYPRNTVLVWRLVAVEENVWIOITD 91		
DB	15	FSGHKVRV-QODPCGGRGLNSKDAGYITSPGYQDYPHQNCSEWVYAPEPNQKIVLNFN 73		
QY	92	ERFLEGDPEDDICKYDEVEVE--EPSDGTILGRWCSGTVPGRQISKGNGRIIRFRFVSDEY 149		
DB	74	PHFLEKHD--CKYDFIETRDGDSADLLGRKCGN-IAPPTIISSGGVLYIKFTSD-Y 128		
QY	150	FPSPGPGFCIHYNI 162		
DB	129	ARQAGGFSRLYEI 141		

RESULT 9

ID	NRPI_MOUSE	STANDARD;	PRT;	923 AA.
AC	P97333;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	NEUROFILIN-1 PRECURSOR (A5 PROTEIN).			
GN	NRPI OR NRP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	{1}			
RN	SEQUENCE FROM N.A.			
RC	SPRATN=BALB/C; TISSUE=Embryonic brain;			
RA	MEDLINE=96353149; PubMed=8748368;			
RX	Kawakami A., Kitsuakawa T., Takagi S., Fujisawa H.;			
RT	"Developmentally regulated expression of a cell surface protein,			
RT	neuropilin, in the mouse nervous system.";			
RL	J. Neurobiol. 29:1-17(1996).			
CC	-1- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE			
CC	CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF			
CC	CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOU			
CC	SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. I			
CC	BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF, THE VEGF-165			
CC	ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN			
CC	INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS			
CC	IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- TISSUE SPECIFICITY: NERVOUS SYSTEM.			
CC	-1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.			
CC	-1- SIMILARITY: CONTAINS 2 CUB DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 1 NAM DOMAIN.			

This SWISS-PROT entry is copyright. It is produced through a collabor
 between the Swiss Institute of Bioinformatics and the EMBL outstat
 the European Bioinformatics Institute. There are no restrictions on
 use by non-profit institutions as long as its content is in no
 modified and this statement is not removed. Usage by and for comm
 entities requires a license agreement (See http://www.isb-sib.ch/annou

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D50086; BAA08789.1; -.
CC MGD; MGI:106206; Nrp.
CC InterPro; IPR000421; -.
CC InterPro; IPR000859; -.
CC InterPro; IPR000998; -.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00629; MAM; 1.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00740; MAM_1; 1.
CC PROSITE; PS00600; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor.
FT SIGNAL. 1 21
FT CHAIN 22 923
FT DOMAIN 22 856
FT TRANSMEM 857 879
FT DOMAIN 880 923
FT DOMAIN 27 141
FT DOMAIN 147 265
FT DOMAIN 275 424
FT DOMAIN 431 583
FT DOMAIN 645 811
FT DISULFID 27 54
FT DISULFID 82 104
FT DISULFID 147 173
FT DISULFID 206 228
FT DISULFID 275 424
FT DISULFID 431 583
FT CARBOHYD 150 150
FT CARBOHYD 261 261
FT CARBOHYD 300 300
FT CARBOHYD 522 522
FT CARBOHYD 842 842
FT SSEQUENCE 923 AA; 103020 MW; 0644B8A170796808 CRC64;
SQ
Query Match 8.68; Score 160.5; DB 1; Length 923;
Best Local Similarity 36.58; Pred. No. 6.1e-06;
Matches 42; Conservativity 16; Mismatches 50; Indels 7; Gaps 4;
QY 50 ITVTSTGSIHSPRPHYPRTVLVRLVAVENNVIOITDERFGLDEPDIDCKYDFV 109
Db 31 IKIENPCYLSPGYPHSPHSEKCEWLQAPQYRIINPHFDLEDRD---CKYDV 87
QY 110 EV--EESDGTILGRWCGSGTVPGKQISGNQIRIRFVSDYFFPSEPFCIHNI 162
Db 88 EVIDGENEGRLMGKFCGK-IAPSPVSSGPFLEFKFVSD-YETHGAGFSIRYEI 140
RESULT 10
NRP2_MOUSE STANDARD; PRT; 931 AA.
AC O35375; O35373; O35374; O35376; O35377; O35378;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165
DE RECEPTOR 2).
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=BALB/C;
RX MEDLINE=97470888; PubMed=9331348;
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
RT "Neurofilin-2, a novel member of the neurofilin family, is a high
```

```
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III.";
RL Neuron 19:547-559(1997).
CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165
CC AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.
CC -!- SUBUNIT: NEUROFILIN-2 PROBABLY FORMS AN HETEROMERIC COMPLEX WITH
CC NEUROFILIN-1 IN ORDER TO BE A FUNCTIONAL SEMAPHORIN E RECEPTOR.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: 6 ISOFORMS; A22 (SHOWN HERE), A0, A5, A17,
CC B0 AND B5; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING CNS, PNS AND IN SOME
CC NONNEURAL TISSUES INCLUDING LIMB BUDS, DEVELOPING BONES, MUSCLES,
CC INTESTINAL EPITHELIUM, KIDNEY, LUNG AND SUBMANDIBULAR GLAND.
CC -!- DEVELOPMENTAL STAGE: THE EXPRESSION PATTERN IS VERY DYNAMIC AND
CC IS DEVELOPMENTALLY REGULATED.
CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF022856; AAC53379.1; -.
CC EMBL; AF022854; AAC53377.1; -.
CC EMBL; AF022855; AAC53378.1; -.
CC EMBL; AF022857; AAC53380.1; -.
CC EMBL; AF022858; AAC53381.1; -.
CC EMBL; AF022861; AAC53382.1; -.
CC MGD; MGI:1100492; Nrp2.
CC InterPro; IPR000421; -.
CC InterPro; IPR000859; -.
CC InterPro; IPR000998; -.
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC PROSITE; PS00629; MAM; 1.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00600; MAM; 2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL. 1 20
FT CHAIN 21 931
FT DOMAIN 21 864
FT TRANSMEM 865 889
FT DOMAIN 890 931
FT DOMAIN 28 142
FT DOMAIN 149 267
FT DOMAIN 277 427
FT DOMAIN 434 592
FT DOMAIN 642 802
FT DOMAIN 838 845
FT DISULFID 28 55
FT DISULFID 83 105
FT DISULFID 149 175
FT DISULFID 208 230
FT DISULFID 277 427
FT DISULFID 434 592
FT CARBOHYD 152 152
FT CARBOHYD 157 157
FT CARBOHYD 629 629
FT CARBOHYD 839 839
FT CARBOHYD 809 813
FT VARSPIC 809 830
FT VARSPIC 814 830
FT VARSPIC 810 931
FT -----
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MISSING (IN ISOFORM A17).
FT MISSING (IN ISOFORM A0).
FT MISSING (IN ISOFORM A5).
FT SSGKESWLYTLDPLITITIAMSSSLGVLGATCAGLLLYCT
```

```
FT FT CTSVGLSSRSCTTLENVNFELDGLKHKVKINHOKCSEA
FT -> GTPPGTPTVDTVPQIPAYVYVMAAGGAVLVLAS
FT VVALVLRHFRFAAKTDRHSITYKTSYTNAGPLAVEPT
FT LTIKLEQERSHC (IN ISOFORM B0).
FT VDIPTETGGEGYEIDDEYEDGWSNSSTSGAGDPSSGK
FT EKSWLYLIDPILITIIANSSGLVLLGATCAGLLIYCTCSYS
FT GLSSRSTTLENVNFELDGLKHKVKINHOKCSEA -> G
FT TLPSPGTPTVDTVPQIPAYVYVMAAGGAVLVLASVVL
FT ALVLRHFRFAAKTDRHSITYKTSYTNAGPLAVEPTITI
FT KLEQERSHC (IN ISOFORM B5).
FT G -> I (IN AAC53380 AND AAC53381).
FT CONFLICT 786 786
FT SEQUENCE 931 AA; 104558 MW; 76F2443F411D2F63 CRC64;
SQ

Query Match 8.6%; Score 159.5; DB 1; Length 931;
Best Local Similarity 33.1%; Pred. No. 7.5e-06;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;

QY 32 FSSNKEONGVODPOHERIITVSTNGSIHSPRPHTYPRNTVLVRLVAEENWITQLTDP 91
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 14 YFSGHEVRSQDPPCGGRPNKSDAGIITSPGYQDPYSHQNCWIVYAPEPNQKIVLNFN 73
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 92 ERFGLEPDDICKYDFEVE--PSPDGTILRCWCGSTVPGKQISKGNIQIRFVSDRY 149
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 74 PHFIEKHD--CRYDFIEIRDGDSADLLGKHGCGN-IAPPTIISGSLVYIKETSD-Y 128
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 150 FPSEPGFCIHVNI 162
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 ARQAGFSLRVEI 141
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
NRPI_HUMAN STANDARD; PRT; 923 AA.
ID NRPI_HUMAN O60461;
AC O14786; O60461;
DT 30-MAY-2000 (Rel. 39, Created)
DI 01-OCT-2000 (Rel. 40, Last sequence update)
DE NEUROPILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
GN NRPI OR NRP OR VEGF165R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
RX MEDLINE=97433084; PubMed=9288753;
RA He Z., Tessier-Lavigne M.;
RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin III".
RL Cell 90:739-751(1997).
RN [2]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
RC TISSUE=Breast;
RX MEDLINE=98188099; PubMed=9529250;
RA Soker S., Takashima S., Mao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";
RL Cell 92:735-745(1998).
RN [3]
RP SEQUENCE FROM N.A. (SOLUBLE/SNRP1 ISOFORM), AND SEQUENCE OF 22-31.
RC TISSUE=Prostatic adenocarcinoma;
RX MEDLINE=20183929; PubMed=10688880;
RA Gagnon M.L., Bieleberg D.R., Gerechtman Z., Miao H.-Q., Takashima S., Soker S., Klagsbrun M.;
RT "Identification of a natural soluble neuropilin-1 that binds vascular endothelial growth factor: In vivo expression and antitumor activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20309748; PubMed=10748121;
```

```
RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid form of vascular endothelial growth factor (VEGF) and of placenta growth factor-2, but only neuropilin-2 functions as a receptor for the 145-amino acid form of VEGF.";
RL J. Biol. Chem. 275:18040-18045(2000).
CC -!- FUNCTION: THE MEMBRANE-BOUND ISOFORM IS A RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. IT BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF. THE VEGF-165 ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS. IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS.
CC -!- FUNCTION: THE SOLUBLE/SNRP1 ISOFORM BINDS VEGF-165 AND APPEARS TO INHIBIT ITS BINDING TO CELLS. IT MAY ALSO INDUCE APOPTOSIS BY SEQUESTERING VEGF-165. MAY BIND AS WELL VARIOUS MEMBERS OF THE SEMAPHORIN FAMILY. ITS EXPRESSION HAS AN AVERSE EFFECT ON BLOOD VESSEL NUMBER AND INTEGRITY.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SOLUBLE/SNRP1 ISOFORM IS SECRETED.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; MEMBRANE-BOUND (SHOWN HERE) AND SOLUBLE/SNRP1; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: THE EXPRESSION OF MEMBRANE-BOUND AND SOLUBLE/SNRP1 ISOFORMS DOES NOT SEEM TO OVERLAP. THE MEMBRANE-BOUND ISOFORM IS EXPRESSED BY THE BLOOD VESSELS OF DIFFERENT TISSUES. IN THE DEVELOPING EMBRYO IT IS FOUND PREDOMINANTLY IN THE NERVOUS SYSTEM. IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN HEART AND PLACENTA; MODERATELY IN LUNG, LIVER, SKELETAL MUSCLE, KIDNEY AND PANCREAS; AND LOW IN ADULT BRAIN. THE SOLUBLE/SNRP1 ISOFORM IS FOUND IN LIVER HEPATOCYTES, KIDNEY DISTAL AND PROXIMAL TUBULES.
CC -!- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF018956; AAC51759.1; -
DR EMBL; AF016050; AAC12921.1; -
DR EMBL; AF145712; AAF44344.1; -
DR MIM; 602069; -
DR InterPro; IPR000421; -
DR InterPro; IPR000859; -
DR InterPro; IPR000998; -
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF00629; MAM; 1.
DR PRINTS; PR00020; MAMDOMAIN.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00600; MAM_2; 1.
DR Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL 1 21
FT CHAIN 22 923
FT DOMAIN 22 856
FT TRANSMEM 857 879
FT DOMAIN 880 923
FT DOMAIN 27 141
FT DOMAIN 147 265
FT DOMAIN 275 424
FT DOMAIN 431 583
FT DOMAIN 645 811
FT NEUROPILIN-1.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT CUB 1.
FT CUB 2.
FT F5/8 TYPE C 1.
FT F5/8 TYPE C 2.
FT MAM.
```

```
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 642 644 EFP->G1K (IN SOLUBLE/SNRP1 ISOFORM).
FT VARSPLIC 645 923 MISSING (IN SOLUBLE/SNRP1 ISOFORM).
FT CONFLICT 26 26 K -> E (IN REF. 1).
FT CONFLICT 749 749 D -> H (IN REF. 2).
FT CONFLICT 855 855 E -> D (IN REF. 2).
SQ SEQUENCE 923 AA; 103120 MW; ADEADC4A849E5D57 CRC64;

Query Match 8.5%; Score 158.5; DB 1; Length 923;
Best Local Similarity 35.7%; Pred. No. 9e-06;
Matches 41; Conservative 18; Mismatches 49; Indels 7; Gaps 4;

Qy 50 IVSTNGSTHSRPHYPRNTVLVRLVAEENWVLIQTFDERFGLDEPDDICKYDFV 109
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 31 IKIESPGYLTSPGYSPHSEKCEWLIQAPQYQIRIMINPHFDLEDRD---CKYDYV 87
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Qy 110 EV--EPPSDGTILGRWCGSGTVPGKQISGNQIRIRFVSDYFPSPGFCIHNYI 162
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 88 EYFDGNGENHGRFGCGK-IAPPPVSSGPFLEIKFVSD-YETHGAGFSIRYEV 140
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 12
NRPL_CHICK STANDARD; PRT; 914 AA.
AC W79795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROPILIN-1 PRECURSOR (A5 PROTEIN).
GN NRPL OR NRP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEHORN; TISSUE=Embryonic brain;
RX MEDLINE=95324761; PubMed=7601310;
RA Takagi S., Kasuya Y., Shimizu M., Matsuura T., Tsuboi M., Kawakami A.,
RA Fujisawa H.,
RT Expression of a cell adhesion molecule, neuropilin, in the
RT developing chick nervous system.*;
RL Dev. Biol. 170:207-222(1995).
CC -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY
CC SIMILARITY). SEEMS TO HAVE CALCIUM-INDEPENDENT CELL ADHESION
CC PROPERTIES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: DEVELOPING NERVOUS SYSTEM; OPTIC TECTUM
CC (LAYERS D AND E OF SGFS), AMACRINE CELLS OF RETINA, NEURITES OF
CC DORSAL ROOT GANGLIA. ALSO EXPRESSED IN NONNEURONAL CELLS, E.G.
CC BLOOD VESSELS IN THE ENTIRE EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 NAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D45416; BAA08256.1; -.
CC InterPro; IPR000421; -.
CC InterPro; IPR000859; -.
CC InterPro; IPR000998; -.
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00754; F5_F8_typeC; 2.
CC Pfam; PF00629; MAM; 1.
CC PRINTS; PR00020; MAMDOMAIN.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00740; MAM_1; 1.
CC PROSITE; PS00060; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Cell adhesion.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 914 NEUROPILIN-1.
FT DOMAIN 20 847 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 848 870 POTENTIAL.
FT DOMAIN 871 914 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 139 CUB 1.
FT DOMAIN 145 263 CUB 2.
FT DOMAIN 273 422 F5/8 TYPE C 1.
FT DOMAIN 429 581 F5/8 TYPE C 2.
FT DOMAIN 636 801 MAM.
FT DISULFID 25 52 PROBABLE.
FT DISULFID 80 102 PROBABLE.
FT DISULFID 145 171 PROBABLE.
FT DISULFID 204 226 PROBABLE.
FT DISULFID 273 422 BY SIMILARITY.
FT DISULFID 429 581 BY SIMILARITY.
SQ SEQUENCE 914 AA; 102480 MW; DD2EE6D6F0CB68C CRC64;

Query Match 8.4%; Score 157; DB 1; Length 914;
Best Local Similarity 29.4%; Pred. No. 1.2e-05;
Matches 52; Conservative 27; Mismatches 68; Indels 30; Gaps 9;

Qy 50 ITVSTNGSIHSRPHYPRNTVLVRLVAEENWVLIQTFDERFGLDEPDDICKYDFV 109
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 29 IKILSPGYLTSPGYSPHSEKCEWLIQAPQYQIRIMINPHFDLEDRD---CKYDYV 85
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Qy 110 EV--EPPSDGTILGRWCGSGTVPGKQISGNQIRIRFVSDYFPSPGFCIHNYI--P 165
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 86 EVIDGDNAEGRWLKCYCGK-IAPPPVSSGPFLEIKFVSD-YETHGAGFSIRYEVFKRGP 143
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Qy 166 Q-----FTEAVSPSLP--PSALPLDL-----LNNATAFTLEDLIRILEPE 206
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 144 ECSRNETSSSGMIKSPGPEKYPNSLCTVIIFAPKMSIILEFESFE-----LEPD 195
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 13
PCOL_HUMAN STANDARD; PRT; 449 AA.
ID PCOL_HUMAN Q15113; O14550;
AC Q15113; O14550;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROCOLLAGEN C-PROTEINASE ENHANCER PROTEIN PRECURSOR (PCPE) (TYPE I
DE DE PROCOLLAGEN COOH-TERMINAL PROTEINASE ENHANCER) (TYPE 1 PROCOLLAGEN C-
DE PROTEINASE ENHANCER PROTEIN).
GN PCOLCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```


RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Boutek J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: REQUIRED FOR NORMAL DORSAL DEVELOPMENT. TLD MAY INTERACT
 CC -1- PHYSICALLY WITH DPP-C PROTEIN.
 CC -1- MISCELLANEOUS: MUTATIONS IN TLD GENE LEAD TO A PARTIAL
 CC TRANSFORMATION OF DORSAL ECTODERM INTO VENTRAL ECTODERM.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
 CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 5 CUB DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; M76976; AAA28491.1; -;
 DR EMBL; U04239; AAC46482.1; -;
 DR EMBL; AE003749; AAF56329.1; -;
 DR PIR; A39288; A39288.
 DR HSP; P00742; IFAX.
 DR MEROPS; M12.010; -;
 DR FlyBase; FBgn0003719; tld.
 DR InterPro; IPR000130; -;
 DR InterPro; IPR000152; -;
 DR InterPro; IPR000561; -;
 DR InterPro; IPR000859; -;
 DR InterPro; IPR001506; -;
 DR InterPro; IPR001881; -;
 DR Pfam; PF01400; Astacin; 1.
 DR Pfam; PF00431; CUB; 5.
 DR Pfam; PF00008; EGF; 2.
 DR PRINTS; PR00480; ASTACIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS01180; CUB; 5.
 DR PROSITE; PS00010; ASX_HYDROXYL; 2.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 2.
 KW Developmental protein; Hydrolase; Metalloprotease; Zinc; Glycoprotein;

KW EGF-like domain; Calcium; Signal; Repeat; Zymogen.
 FT SIGNAL 1 27 POTENTIAL.
 FT PROPEP 28 126
 FT CHAIN 127 1057 DORSAL-VENTRAL PATTERNING TOLLOID
 FT PROTEIN DORSAL-VENTRAL PATTERNING TOLLOID
 FT DOMAIN 127 329 METALLOPROTEASE.
 FT DOMAIN 330 467 CUB.
 FT DOMAIN 468 580 CUB.
 FT DOMAIN 581 621 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 624 742 CUB.
 FT DOMAIN 743 783 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 787 899 CUB.
 FT DOMAIN 900 1016 CUB.
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
 FT SITE 233 237 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 315 317 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 330 380 BY SIMILARITY.
 FT DISULFID 407 429 BY SIMILARITY.
 FT DISULFID 468 495 BY SIMILARITY.
 FT DISULFID 522 544 BY SIMILARITY.
 FT DISULFID 585 596 BY SIMILARITY.
 FT DISULFID 592 605 BY SIMILARITY.
 FT DISULFID 607 620 BY SIMILARITY.
 FT DISULFID 624 652 BY SIMILARITY.
 FT DISULFID 683 706 BY SIMILARITY.
 FT DISULFID 747 758 BY SIMILARITY.
 FT DISULFID 754 767 BY SIMILARITY.
 FT DISULFID 787 813 BY SIMILARITY.
 FT DISULFID 840 862 BY SIMILARITY.
 FT DISULFID 900 930 BY SIMILARITY.
 FT DISULFID 957 979 BY SIMILARITY.
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 533 533 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 667 667 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 781 781 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 854 854 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 908 908 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1057 AA; 120575 MW; 76F4B5AEB7996FEA CRC64;
 Query Match 8.0%; Score 148.5; DB 1; Length 1057;
 Best Local Similarity 38.0%; Pred. No. 7.7e-05;
 Matches 38; Conservative 20; Mismatches 35; Indels 7; Gaps 5;
 QY 50 ITVSTNGSIHSRPHPTYPRTNTVLVRLVAVENWVQLTFDERFGLDEPDDICKYDFV 109
 DB 472 LKLTQDSIDSPNYPMDYMPDKECVWRITA-PDNHQVALKF-QSFELE--KHDGCAYDFV 527
 QY 110 EVEEE--PSDGTILGRWCGSGTVPKQISKGNQIRIRFVSD 147
 DB 528 EIRDGNHSDSRLLIGRCGDKLPNPK-TRSNQMYIRFVSD 566
 RESULT 15
 ID MAS2_HUMAN STANDARD; PRT; 686 AA.
 AC 000187; O75754;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE MANNAN-BINDING LECTIN SERINE PROTEASE 2 PRECURSOR (EC 3.4.21.-)
 DE (MANNOSE-BINDING PROTEIN ASSOCIATED SERINE PROTEASE 2) (MASP-2).
 GN MASP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:07:48 ; Search time 27.91 Seconds
(without alignments)
1635.444 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 1858
Sequence: 1 MSFLGLLLTALAGROGT.....DVALEHHEECDCVCRGSTGG 345

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1858	100.0	345	4	Q9UL22	Q9ul22 homo sapien
2	1851	99.6	345	4	Q9NRA1	Q9nra1 homo sapien
3	1667	89.7	345	11	Q9QY71	Q9qy71 mus musculu
4	1649	88.2	345	11	Q9BQX6	Q9eqx6 rattus norv
5	1638	88.2	345	11	Q9JHV8	Q9jvh8 mus musculu
6	1617	87.0	345	13	Q9I946	Q9i946 gallus gall
7	742.5	40.0	370	4	Q9GZP0	Q9gzp0 homo sapien
8	736	39.6	370	11	Q9BQT1	Q9eqt1 rattus norv
9	183.5	9.9	3623	4	O60494	O60494 homo sapien
10	182	9.8	691	13	O57658	O57658 gallus gall
11	176	9.5	977	13	Q91925	Q91925 xenopus lae
12	172	9.3	1012	11	Q9WVM6	Q9wvm6 mus musculu
13	171	9.2	735	13	O57381	O57381 xenopus lae
14	171	9.2	926	4	Q9UQ00	Q9uq00 homo sapien
15	171	9.2	1015	4	Q9Y6L7	Q9y6l7 homo sapien
16	171	9.2	3623	11	O70244	O70244 rattus norv
17	169	9.1	241	11	O92135	O9z135 rattus norv
18	164	8.8	3620	6	Q9TU53	Q9tu53 canis fami
19	163.5	8.8	555	4	Q9H2E2	Q9h2e2 homo sapien

20	163.5	8.8	901	4	Q9H2E4	Q9h2e4 homo sapien
21	163.5	8.8	901	4	Q9H2D5	Q9h2d5 homo sapien
22	163.5	8.8	906	4	Q9H2E3	Q9h2e3 homo sapien
23	163.5	8.8	906	4	Q9H2D4	Q9h2d4 homo sapien
24	163.5	8.8	921	11	Q9QX38	Q9qxx38 rattus norv
25	162.5	8.7	326	11	O35251	O35251 rattus norv
26	159	8.6	1022	13	O57460	O57460 brachydanio
27	158.5	8.5	704	4	Q9H2E1	Q9h2e1 homo sapien
28	156	8.4	1008	13	Q9DER7	Q9der7 gallus gall
29	153	8.2	415	4	Q9UKZ9	Q9ukz9 homo sapien
30	153	8.2	1019	13	O57382	O57382 xenopus lae
31	149	8.0	1013	4	O43897	O43897 homo sapien
32	149	8.0	1013	4	Q9NQS4	Q9nqs4 homo sapien
33	148	8.0	358	13	O57434	O57434 fugu rubrip
34	148	8.0	423	13	O57433	O57433 fugu rubrip
35	148	8.0	1013	11	O62381	O62381 mus musculu
36	147	7.9	354	4	O43915	O43915 homo sapien
37	146	7.9	746	5	O01654	O01654 halocynthia
38	143.5	7.7	686	4	Q9Y270	Q9y270 homo sapien
39	142.5	7.7	212	11	O09020	O09020 rattus norv
40	142.5	7.7	701	11	O9JJS9	O9jjs9 rattus norv
41	140	7.5	358	11	P97946	P97946 mus musculu
42	139.5	7.5	1070	5	P91972	P91972 aplysia cal
43	138.5	7.5	597	5	Q26051	Q26051 paracentrot
44	137.5	7.4	699	4	O95570	O95570 homo sapien
45	137.5	7.4	699	4	Q9UF09	Q9uf09 homo sapien

ALIGNMENTS

RESULT 1
Q9UL22
ID Q9UL22 PRELIMINARY; PRT: 345 AA.
AC Q9UL22;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE SECRETORY GROWTH FACTOR-LIKE PROTEIN FALLOTEIN (SPINAL CORD-DERIVED GROWTH FACTOR).
GN HSCDGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBIFaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Tsai Y.J., Lee R.K.K., Lin S.P.;
RT "Fallotein, a novel growth factor like gene identified in human uterus.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=20317014; PubMed=10858496;
RA Hamada T., Oi-Tel K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDF, is a unique member of the PDGF/VEGF family.";
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AF091434; AAF00049.1; -;
DR EMBL; AB033831; BAB03286.1; -;
DR InterPro; IPR000072; -;
DR InterPro; IPR000859; -;
DR Pfam; PF00341; PDGF; 1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS02078; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 39029 MW; CDE9E51F40633E78 CRC64;

Query Match 100.0%; Score 1858; DB 4; Length 345;

```
Best Local Similarity 100.0%; Pred. No. 2e-165;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFGLLLTTSALAGORCTQAESNLSSKQFSSNKEQNGVQDPOHERIITVSTNGSIHS 60
Db 1 MSFGLLLTTSALAGORCTQAESNLSSKQFSSNKEQNGVQDPOHERIITVSTNGSIHS 60
QY 61 PRFHTYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
Db 61 PRFHTYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
QY 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
Db 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
QY 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVFGKSRVVDNL 240
Db 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVFGKSRVVDNL 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
QY 301 VTKKYHEVLQRLPKTGVRLHKSLLTDVALEHHEECDCVCRGSTG 345
Db 301 VTKKYHEVLQRLPKTGVRLHKSLLTDVALEHHEECDCVCRGSTG 345

RESULT 2
Q9NR1 PRELIMINARY; PRT; 345 AA.
ID AC Q9NR1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Li X., Ponten A., Aase K., Karlsson L., Abramsson A., Uutela M.,
RA Backstrom G., Hellstrom M., Bostrom H., Li H., Soriano P.,
RA Betsholtz C., Heldin C.-H., Alitalo K., Ostman A., Eriksson U.;
RT "PDGF-C is a novel protease-activated ligand for the PDGF alpha
RT receptor.";
RL Nat. Cell Biol. 0:0-0(2000).
DR EMBL; AF244813; AAF80597.1; -
DR InterPro; IPR000072; -
DR InterPro; IPR000859; -
DR Pfam; PF00341; PDGF_1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 35043 MW; 590889CEA55C5EA CRC64;

Query Match 99.6%; Score 1851; DB 4; Length 345;
Best Local Similarity 99.4%; Pred. No. 9.1e-165;
Matches 343; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFGLLLTTSALAGORCTQAESNLSSKQFSSNKEQNGVQDPOHERIITVSTNGSIHS 60
Db 1 MSFGLLLTTSALAGORCTQAESNLSSKQFSSNKEQNGVQDPOHERIITVSTNGSIHS 60
QY 61 PRFHTYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
Db 61 PRFHTYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
QY 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
Db 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
```

```
Db 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
QY 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVFGKSRVVDNL 240
Db 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVFGKSRVVDNL 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
QY 301 VTKKYHEVLQRLPKTGVRLHKSLLTDVALEHHEECDCVCRGSTG 345
Db 301 VTKKYHEVLQRLPKTGVRLHKSLLTDVALEHHEECDCVCRGSTG 345

RESULT 3
Q9QY71 PRELIMINARY; PRT; 345 AA.
ID AC Q9QY71;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-WAR-2001 (Tremblrel. 16, Last annotation update)
DE FALLOTEIN.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Tsai Y.-J., Lee R.K.-K., Chen Y.-H., Lin S.-P., Cheng W.T.-K.;
RT "cDNA cloning of follotoin from mouse ovary.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117608; AAF22516.1; -
DR InterPro; IPR000072; -
DR InterPro; IPR000859; -
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38741 MW; 3A58A1F701B84EA2 CRC64;

Query Match 89.7%; Score 1667; DB 11; Length 345;
Best Local Similarity 87.0%; Pred. No. 1.4e-147;
Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MSFGLLLTTSALAGORCTQAESNLSSKQFSSNKEQNGVQDPOHERIITVSTNGSIHS 60
Db 1 MSFGLLLTTSALAGORCTQAESNLSSKQFSSNKEQNGVQDPOHERIITVSTNGSIHS 60
QY 61 PRFHTYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
Db 61 PRFHTYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
QY 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
Db 121 GRWCGSTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPVLPPSA 180
QY 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVFGKSRVVDNL 240
Db 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLKAFVFGKSRVVDNL 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
QY 301 VTKKYHEVLQRLPKTGVRLHKSLLTDVALEHHEECDCVCRGSTG 345
Db 301 VTKKYHEVLQRLPKTGVRLHKSLLTDVALEHHEECDCVCRGSTG 345
```

```
RESULT 4
Q9EQX6 PRELIMINARY; PRT; 345 AA.
AC Q9EQX6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN RSCDGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NISTAR: TISSUE=KIDNEY;
RA Hamada T., Ui-Tei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDF-B, a Novel Growth Factor Homologous to
RT SCDF/PDGF-C/fallotelin.";
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
DR EMBL; AB033830; BAB19969.1; -.
SQ SEQUENCE 345 AA; 38734 MW; F296DA6E9B765D10 CRC64;

Query Match 88.8%; Score 1649; DB 11; Length 345;
Best Local Similarity 85.8%; Pred. No. 6.8e-146;
Matches 296; Conservative 29; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGORQGTQAESNLSSKQFSSNKQNGVODPQHERITVTNGSIHS 60
DB 1 MLLGLLLTSALAGORTGTAEENSLSSKQLSSKQNGVODPQHERITVTNGSIHS 60
QY 61 PRPHPTYPNTVLRVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
DB 61 PRPHPTYPNTVLRVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
QY 61 PRPHPTYPNTVLRVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
DB 61 PRPHPTYPNTVLRVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
QY 121 GRWCGSTVPKGKQISGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPSVLPSSA 180
DB 121 GRWCGSTVPKGKQISGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPSVLPSSA 180
QY 181 LPDLNNATVAFSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 LPDLNNATVAFSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
QY 241 LEEVRLYSCPTPNFSVSIREELKRTDTIFWPGCLLYKRCGNCACCLHNCNECQCVPSK 300
DB 241 LKEEVKLYSCPTPNFSVSIREELKRTDTIFWPGCLLYKRCGNCACCLHNCNECQCVPSK 300
QY 301 VTKKYHEVLQRPKTVGRGLHKLSTDALEHHECDVCVCGSTGG 345
DB 301 VTKKYHEVLQRPKTVGRGLHKLSTDALEHHECDVCVCGSTGG 345

RESULT 5
Q9JHV8 PRELIMINARY; PRT; 345 AA.
AC Q9JHV8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
GN PDGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS-WEBSTER/NIH;
RA Ding H., Wu X., Kim I., Tam P.P.L., Koh G.Y., Nagy A.;
RT "The mouse pdgfc gene: Dynamic expression in embryonic tissues during
RT organogenesis.";
RL Mech. Dev. 0:0-0(2000).
```

```
DR EMBL; AF286725; AAF91483.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38886 MW; FA1486BED6D362F8 CRC64;

Query Match 88.2%; Score 1638; DB 11; Length 345;
Best Local Similarity 85.8%; Pred. No. 7.2e-145;
Matches 296; Conservative 27; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGORQGTQAESNLSSKQFSSNKQNGVODPQHERITVTNGSIHS 60
DB 1 MLLGLLLTSALAGORTGTAEENSLSSKQLSSKQNGVODPQHERITVTNGSIHS 60
QY 61 PRPHPTYPNTVLRVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
DB 61 PRPHPTYPNTVLRVAVENWVQLTDFDERFGLDEDDICKYDFVEVEEPSDGTIL 120
QY 121 GRWCGSTVPKGKQISGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPSVLPSSA 180
DB 121 GRWCGSTVPKGKQISGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPSVLPSSA 180
QY 181 LPDLNNATVAFSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 LPDLNNATVAFSTLEDLIRYLEPERWQDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
QY 241 LEEVRLYSCPTPNFSVSIREELKRTDTIFWPGCLLYKRCGNCACCLHNCNECQCVPSK 300
DB 241 LKEEVKLYSCPTPNFSVSIREELKRTDTIFWPGCLLYKRCGNCACCLHNCNECQCVPSK 300
QY 301 VTKKYHEVLQRPKTVGRGLHKLSTDALEHHECDVCVCGSTGG 345
DB 301 VTKKYHEVLQRPKTVGRGLHKLSTDALEHHECDVCVCGSTGG 345

RESULT 6
Q91946 PRELIMINARY; PRT; 345 AA.
AC Q91946;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN SCDFG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=SPINAL CORD;
RA MEDLINE-20317014; PubMed-10858496;
RA Hamada T., Ui-Tei K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDFG, is a unique
RT member of the PDGF/VEGF family.";
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AB033829; BAB03265.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38940 MW; 97ACEA992BF5128C CRC64;

Query Match 87.0%; Score 1617; DB 13; Length 345;
Best Local Similarity 84.1%; Pred. No. 6.6e-143;
```



```
DR InterPro: IPR000130; -
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR InterPro: IPR000859; -
DR InterPro: IPR001506; -
DR InterPro: IPR001881; -
DR Pfam: PF000008; EGF; 2.
DR Pfam: PF00431; CUB; 5.
DR PRINTS: PR00480; ASTACIN.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS01180; CUB; 5.
DR PROSITE: PS01186; EGF; 2.
DR PROSITE: PS01187; EGF CA; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART: SW00042; CUB; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 1012 AA; 113252 MW; 72EE268A4D8C5FE CRC64;
```

Query Match 9.3%; Score 172; DB 11; Length 1012;
Best Local Similarity 41.3%; Pred.No. 2.8e-07;
Matches 45; Conservative 17; Mismatches 37; Indels 10; Gaps

```
DQ   55 NGSIHSPRPHTVPRTVLVWRLLVA--VEENVVTQLTFDERFGLGEDDDICKYDFVEEVE 113  
    ||||| : | ||| ||||| : : | ||| : ||||| |||||  
DB   625 NGITISFGWPKPEPTNKNCWQVAVPYQRISLP---EAFELEG--NDVKYDFEVRS 679  
            : : ||||| : | : ||| : ||| : ||| :  
DQ   114 --PSDTILGRWCSTGTPGKIQISKGNQIRIRFVSDEYPSPGPFIHY 160  
            : : ||||| : | : ||| : ||| : ||| :  
DB   680 GLSDAKLHGKFCGSET-PEVITQSNNMRVFESKDNTV-SKRGFRAHF 726
```

RESULT 13

ID	OS7381
IC	PRELIMINARY; PRT; 735 AA.
AD	O57381;
OT	O57381;
DT	01-JUN-1998 (TrEMBLrel. 06, Created)
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DN	BONE MORPHOGENETIC PROTEIN 1B.
GE	BMP-1.
GC	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidea;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8355;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-Ovary;
RA	Goodman S., Albano R., Matthews G., Tannahill D., Dale L.;
RL	Dev. Biol. 0:0-0(0);
DR	EMBL; Y09660; CAA70853.1; -.
DR	HSP; P00736; IAPQ.
DR	InterPro: IPR000130; -.
DR	InterPro: IPR000152; -.
DR	InterPro: IPR000561; -.
DR	InterPro: IPR000859; -.
DR	InterPro: IPR001506; -.
DR	InterPro: IPR001881; -.
DR	Pfam: PF00008; EGF; 1.
DR	Pfam: PF00431; CUB; 3.
DR	Pfam: PF01400; Astacin; 1.
DR	PRINTS: PR00480; ASTACIN.
DR	PROSITE: PS00010; ASX_HYDROXYL; 1.
DR	PROSITE: PS01180; CUB; 3.
DR	PROSITE: PS01186; EGF; 1.
DR	PROSITE: PS01187; EGF CA; 1.
DR	PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR	SMART; SW00042; CUB; 1.
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ	SEQUENCE 735 AA; 83575 MW; 45B29C813f79DBE2 CRC64;

```
Query Match          9.2%; Score 171; DB 13; Length 735;
Best Local Similarity 43.3%; Pred. No. 2.2e-07;
Matches 45; Conservative 14; Mismatches 37; Indels 8; Gaps 5;

QY 55 NGSIHSPRPHTYPRNTVLVRLVAEENWVWVQIQLTDEREGLEDPEDDICKYDFVEVEE- 113
DB 589 NGITSPGWPKEYPNTKNCVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 644

QY 114 -PSDGTILGRWCSTGTPVKQISKGNQIRIRFVSDEYFPSPGFCIH 156
DB 645 LRSDSLKHGKFCGT-ELPAVITSOYNNMRIEFKSDNTV-SKKGF 686

RESULT 14
ID Q9UQ00 PRELIMINARY; PRT; 926 AA.
AC Q9UQ00;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE KIAA0932 PROTEIN (FRAGMENT).
GN KIAA0932.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirotsawa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:63-70(1999).
DR EMBL; AB023149; BAA76776.1; -.
DR HSSP; P00736; IAPQ.
DR InterPro; IPR000130; -.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000859; -.
DR InterPro; IPR001506; -.
DR InterPro; IPR001881; -.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF01400; Astacin; 1.
DR PRINTS; PR00480; ASTACIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01180; CUB; 5.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00042; CUB; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER
SQ SEQUENCE 926 AA; 103873 MW; 58FA35CDDDE10970B CRC64;
```

```
Query Match          9.2%; Score 171; DB 4; Length 926;
Best Local Similarity 43.5%; Pred. No. 3e-07;
Matches 47; Conservative 14; Mismatches 39; Indels 8; Gaps 6;

QY 55 NGSIHSPRPHTYPRNTVLVRLVAEENWVWVQIQLTDEREGLEDPEDDICKYDFVEVEE- 113
DB 539 NGITSPGWPKEYPNTKNCVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 594

QY 114 -PSDGTILGRWCSTGTPVKQISKGNQIRIRFVSDEYFPSPGFCIH 160
DB 595 LSPDAKLHGRFCGSET-PEVITSQSNMRVFEFKSDNTV-SKRGFRAHF 640
```

```
RESULT 15
QY 09Y6L7 PRELIMINARY; PRT; 1015 AA.
AC 09Y6L7;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE TOLLOID-LIKE 2 PROTEIN.
GN TLL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Scott I.C., Greenspan D.S.;
RT "Sequence of the human mammalian tolloid-like 2 (mTll-2) and
RT chromosomal localisation of the cognate gene TLL2.";
RL Dev. Biol. 0:0-0(1999).
DR EMBL; AF059516; AAD42979.1; -.
DR HSSP; P00736; IAPQ.
DR InterPro; IPR000130; -.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000859; -.
DR InterPro; IPR001506; -.
DR InterPro; IPR001881; -.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00431; CUB; 5.
DR Pfam; PF01400; Astacin; 1.
DR PRINTS; PR00480; ASTACIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01180; CUB; 5.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00042; CUB; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 1015 AA; 113536 MW; 25F5B23065861593 CRC64;
```

```
Query Match          9.2%; Score 171; DB 4; Length 1015;
Best Local Similarity 43.5%; Pred. No. 3.4e-07;
Matches 47; Conservative 14; Mismatches 39; Indels 8; Gaps 6;

QY 55 NGSIHSPRPHTYPRNTVLVRLVAEENWVWVQIQLTDEREGLEDPEDDICKYDFVEVEE- 113
DB 628 NGITSPGWPKEYPNTKNCVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 683

QY 114 -PSDGTILGRWCSTGTPVKQISKGNQIRIRFVSDEYFPSPGFCIH 160
DB 684 LSPDAKLHGRFCGSET-PEVITSQSNMRVFEFKSDNTV-SKRGFRAHF 729
```

Search completed: June 29, 2001, 17:11:34
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 00:39:28 ; Search time 2381.26 Seconds
(without alignments)
11458.266 Million cell updates/sec

Title: US-09-457-066-1
Perfect score: 1764
Sequence: 1 attatgtgaaactaccctg.....cttgtgtgtgtgatagga 1764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: gb_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

- em_ov.*
44: em_pat.*
45: em_ph.*
46: em_pl.*
47: em_pl.*
48: em_ro.*
49: em_ro.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil.*
60: gb_vil.*
61: gb_vil.*
62: gb_vil.*
63: gb_vil.*
64: gb_vil.*
65: gb_vil.*
66: gb_vil.*
67: gb_vil.*
68: gb_vil.*
69: gb_vil.*
70: gb_vil.*
71: gb_vil.*
72: gb_vil.*
73: gb_vil.*
74: gb_vil.*
75: gb_vil.*
76: gb_vil.*
77: gb_vil.*
78: gb_vil.*
79: gb_vil.*
80: gb_vil.*
81: gb_vil.*
82: gb_vil.*
83: gb_vil.*
84: gb_vil.*
85: gb_vil.*
86: gb_vil.*
87: gb_vil.*
88: gb_vil.*
89: gb_vil.*
90: gb_vil.*
91: gb_vil.*
92: gb_vil.*
93: gb_vil.*
94: gb_vil.*
95: gb_vil.*
96: gb_vil.*
97: gb_vil.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1746	99.0	1760	9	AX044518	AX044518 Sequence
2	1734	98.3	3007	88	AF091434	AF091434 Homo sapi
3	1732.4	98.2	2849	9	AX047650	AX047650 Sequence
4	1701	96.4	2152	89	AF244813	AF244813 Homo sapi
5	1574	89.2	1817	85	AB033831	AB033831 Homo sapi
6	1035	58.7	1035	9	AX027935	AX027935 Sequence
7	1035	58.7	1035	91	AX028032	AX028032 Sequence
8	883.2	50.1	3571	9	AX044520	AX044520 Sequence

9 881.6 50.0 2692 94 AF117608 Mus muscu
10 780 44.2 1116 94 AB033830 Rattus no
11 779.8 44.2 1675 8 AB033829 Gallus ga
12 773 43.8 1038 94 AF286725 Mus muscu
13 662.4 37.6 191314 61 AC009582 Homo sapi
14 447 25.3 504 9 AX027960 Sequence
15 447 25.3 504 91 AX028057 Sequence
16 317.8 18.0 82113 63 AC015451 Homo sapi
17 283.4 16.1 289 9 AX027970 Sequence
18 283.4 16.1 289 9 AX027993 Sequence
19 283.4 16.1 289 56 AX028067 Sequence
20 283.4 16.1 289 56 AX028090 Sequence
21 279 15.8 279 9 AX027968 Sequence
22 279 15.8 279 9 AX027989 Sequence
23 279 15.8 279 56 AX028065 Sequence
24 279 15.8 279 56 AX028086 Sequence
25 278.6 15.8 284 9 AX027963 Sequence
26 278.6 15.8 284 9 AX027982 Sequence
27 278.6 15.8 284 56 AX028060 Sequence
28 278.6 15.8 284 56 AX028079 Sequence
29 273 15.5 300 9 AX027962 Sequence
30 273 15.5 300 9 AX027981 Sequence
31 273 15.5 300 56 AX028059 Sequence
32 273 15.5 300 56 AX028078 Sequence
33 272.4 15.4 152779 64 AC015837 Homo sapi
34 261 14.8 261 9 AX027967 Sequence
35 261 14.8 261 9 AX027988 Sequence
36 261 14.8 261 56 AX028064 Sequence
37 261 14.8 261 56 AX028085 Sequence
38 259.4 14.7 275 9 AX027964 Sequence
39 259.4 14.7 275 9 AX027984 Sequence
40 259.4 14.7 275 56 AX028061 Sequence
41 259.4 14.7 275 56 AX028081 Sequence
42 255 14.5 278 9 AX027965 Sequence
43 255 14.5 278 9 AX027986 Sequence
44 255 14.5 278 56 AX028062 Sequence
45 255 14.5 278 56 AX028083 Sequence

ALIGNMENTS

RESULT 1
AX044518
LOCUS AX044518 1760 bp DNA PAT 24-NOV-2000
DEFINITION Sequence 32 from Patent WO00066736.
ACCESSION AX044518
VERSION AX044518.1 GI:11343373
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1760)
AUTHORS Gilbert, T., Hart, C.E., Sheppard, P.O. and Gilbertson, D.G.
TITLE Growth factor homolog zveg4
JOURNAL Patent: WO 0066736-A 32 09-NOV-2000;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..1760
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1191
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC11710.1"
/db_xref="GI:11343374"
/translation="MSLFGILLTTSALAQROGTOAESNLSSKFORSSNKQGVQDP
QHERLITVSTNGSTHSRPHPTYPNTVLVRLVAEENVIQLTDFERFGLDEPDD
ICKYDFVEVPSDGTILGRNCGSGTVPGKQISKQNIIRIRVSPSEPGFCHY
NIVMPQTEAVSPVLPSPALPDLLNNATFSTLEDLIRYLEPFWOLDLEDIRP
TWQLGKAFVGRKSRVVDLNLTEEVRLYSCTPRNFVSIRREELKSLDTIFWPGCLL
VKRGGNCACCLHNCQCQVPSKVKRYKHYEVLQRLPRKTGVRGLHSLTDVALEHHEE

QY 961 ttctgcccagggttgctctctcgttaaacgctgtggtggtggaactgtgctgttctctccac 1020
Db 961 TTCTGGCCAGGTTGCTCTCCCTGGTTAAACGCTGTGGTGGAACTGTGCTGTGCTCCAC 1020
QY 1021 aattgcaatgaatgcaatgtgcccaagcaaatgactactaaaaataaccagaggtcctt 1080
Db 1021 AATTGCAATGAATGTCAATGTGCCCAAGCAAGTTACTAAAAAATACCAGAGGTCCTT 1080
QY 1081 cagttgagaccacaagaccggtgtcaggggattgcacaaatacactcaccgagctggccctg 1140
Db 1081 CAGTTGAGACCACAAGACCAGGCTGTGAGGGATTGCACAAATCACTCACCGAGTGCCCTG 1140
QY 1141 gagcacatgagagtgctgactgtgtgagagggagcagagagatagccgcatac 1200
Db 1141 GAGCACATGAGAGTGTGACTGTGTGAGAGGGAGCAGAGAGATAGCCGCATCAC 1200
QY 1201 caccagcactctgccagagctgtcagtllyllycagaggtgacttctattagagaac 1260
Db 1201 CACCAGCAGCTGTGCCAGAGCTGTGCAGT- ---GCAGTGGCTGATTCATTATAGAGAAC 1260
QY 1261 gtatgcttatctccatcccttaactcagttgttctcagagacctttcatcttcagg 1320
Db 1261 GTATGCTTATCTCCATCTTAATCTCAGTTGTTGTCTCAAGGACCTTTCATCTTCAGG 1320
QY 1321 attacagtgcatctgaaagagagagacatacaacagaattagagtggtgcaacagctc 1380
Db 1321 AATTACAGTGCAATCTCAAAAGAGAGAGATCAAAACAGAAATTAGGAGTTGTCAACAGCTC 1380
QY 1381 ttttgagagggcctaaagacagaggaagaaaggctctcaatcgtgaaagaaaaataaa 1440
Db 1381 TTTTGAGAGGGCCTAAAGACAGAGGAGAAAGGCTCTCAATCTGTGAAAGAAAAATAAA 1440
QY 1441 tgtgtatataatagatcacacagctagttcagagttaccatgtacatgtattccactagct 1500
Db 1441 TGTGTATATAATAGATCACACAGCTAGTTTTCAGAGTTACCATGTACGATTTCCATAGCT 1500
QY 1501 ggggtctgtatttcagttcttcgatacagcgttagggtaattgctacagacagaaagaaac 1560
Db 1501 GGGTCTGTATTTTTCAGTTCTTCGATACGCTTAGGGTAAATGTGTCAGTACAGGAAAAAAC 1560
QY 1561 tbtgcaagtgtgacacctgtatccgtgtcctgtcttaactctaaagctccatgtcctgggc 1620
Db 1561 TGTGCAAGTGTGACACCTGATTCGCTTGCCTTGTCTTAACTCTTAAAGCTCCATGTCTGGC 1620
QY 1621 ctataatctataaactgtgatt 1680
Db 1621 CTATAATCTGTATATAAATCTGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTCTCATAT 1680
QY 1681 accagaaactctatgtactacaaacacctgggttttttaaaaggaaactatgttgcctatgaat 1740
Db 1681 ACCAGAACTCTATGTACTACAAACCTGGTTTTTAAAGGAACATGTGTGCTATGAAT 1740
QY 1741 taaactgtgtcgtgtgatagga 1764
Db 1741 TAAACTGTGTGCTGTGTATAGGA 1760

RESULT 2
AF091434 3007 bp mRNA PRI 22-JUN-2000
LOCUS Homo sapiens secretory growth factor-like protein fallotein mRNA, complete cds.
DEFINITION
ACCESSION AF091434
VERSION AF091434.1 GI:6002592
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3007)
AUTHORS Tsai, Y., Lee, R. K., Lin, S. and Chen, Y.
TITLE Identification of a novel platelet-derived growth factor-like gene, fallotein, in the human reproductive tract

JOURNAL Biochim. Biophys. Acta 1492 (1), 196-202 (2000)
MEDLINE 20461776
REFERENCE 2 (bases 1 to 3007)
AUTHORS Tsai, Y., Lee, R. K. K. and Lin, S. P.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-1998) Dept. Medical Research, Mackay Memorial Hospital, 45 Min Sheng Road, Tamshui, Taipei County 25115, Taiwan
FEATURES
source Location/Qualifiers
1. 3007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="uterus"
492..1529
/evidence="not_experimental"
/product="secretory growth factor-like protein fallotein"
/protein_id="AAF00049.1"
/db_xref="GI:6002593"
/translation="MSLFGLLLLLSALAGRQGTQAESNLSSKFQSSNKQNGVQDP
QHERITVSTNGSIHSPRPHTYPRNTVLVWLVAVEENWVLIQTFDFRFGLEDPEDD
ICKYDFEVEBPSTGTILGRWCGSGTVPKQISKGNOIRFVSDVFPSPGFCIH
NIVMPQTEAVSPSVLPSPALPLDILNNAITAFSTLEDLIRYLPERWQLEDLYRP
TWQLGKAFVGRKSRVVLDNLLEEVRLISCTPNFVSFIREKRTDTTFWPGCLL
VKRCGNCACCLHNCBQCQVPSKVTKKYHEVLQRLPKRTGYRGLHLSLTDVALEHHE
CDCVCRGSTGG"
polyA_signal 858 a 613 c 635 g 901 t
BASE COUNT 2974..2979
ORIGIN
Query Match 98.3%; Score 1734; DB 88; Length 3007;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1759; Conservative 0; Mismatches 0; Indels 5; Gaps 2;
QY 1 attatgtgaaactaccctggattctctgtgccagagcaggtcggcgtccacccc 60
Db 339 ATTATGTGAAACTACCCTGCGATTCTCTGTGCCAGACAGGCTCGGGCTTCCACCCC 398
QY 61 agtgagccttccctggcgtgtgaaagagactcggagtcgctgtccaaagtgc 120
Db 399 AGTGAGCCTTCCCTGGCGTGGTGAAGAGACTCGGGAGTCGGTCTTCCAAAGTGCC 458
QY 121 cgcctgtgagtgctctcaccagctcagcacaatgagcctcttcgggcttctcctgctg 180
Db 459 CGCGGTGAGTGAGCTCTCACCCAGTCAGCCAAATGAGCCTTTCGGGCTTCTCTGCTG 518
QY 181 acatgcctcggcggccagagacagagggactcaggcggaatccaacctgagtagtaaa 240
Db 519 ACATCTGCCCTGGCGGCCAGAGACAGGGGACTCAGGGGGAATCCAACCTGAGTAGTAA 578
QY 241 ttcaagtttccagcaacaagaaacagagagagagagagagagagagagagagagagag 300
Db 579 TTCCAGTTTCCAGCAACAAGAACAGACGAGGATACAGATCCTCAGCATGAGAGAAT 638
QY 301 attactgtctactaatggaagtattcacagcccaagtttccctcactactatccaaga 360
Db 639 ATTACTGTCTACTTAATGGAAGTATTACAGCCCAAGGTTTCCCTACTACTTATCCAGA 698
QY 361 aatacggcttggtgtagagattagtagcagtagaggaataatgtaggatacaacttacg 420
Db 699 AATACGGCTTGGTATGGAGATTAGTAGCAGTAGAGGAAATGTATGGATACACTTACG 758
QY 421 ttgtatgaagatttgggttgaagaccagagagatgacatatgcaagtatgattttgta 480
Db 759 TTTGATGAAGATTTGGGCTTGAAGACCCAGAAAGATGACATATGCAAGTATGATTTTCTA 818
QY 481 gaagttaggaaccagtgatggaaactatatagggcgtcgtggtggttctggtactgta 540
Db 819 GAACTTGAGAACCCAGTATGGAACATATATAGGGCCCTGGTGTGGTCTGGTACTGTA 878
QY 541 ccaggaacacagattctcacaaggaaatacaattaggaataagattgtctctgatgaat 600
Db 879 CCAGGAAACAGATTTCTAAAGGAATCAATATGAGTAAGATTTGTATCTGTGATGAAT 938

Qy 601 ttctcttgacacagggttctgcatccactacatacaacattgtcatgccacaattcacagaa 660
Db 939 TTTCTCTCTGAACACAGGGTTCTGCATCCACTACAACTTGTATGCCACAAATTCAGAA 998
Qy 661 gctgtgagctctcagtgctaccccccttcagcttggcaactgttggcaagcctgttcaatagct 720
Db 999 GCTGTGAGTCTCTCAGTGTCTACCTCTTCCAGCTTTGCGCACTGGACCTGCTTAATAGCT 1058
Qy 721 ataactgcttagtaccttgagagaccttattcgatatcttgaaaccagagagatggcag 780
Db 1059 ATAAGTGGCTTTAGTACCTTGGAGACCTTATTCGATATCTTGAACAGAGAGATGGCAG 1118
Qy 781 ttggacttagaagatctatataggccaaacttggcaacttcttggcaagccttcttggcttt 840
Db 1119 TTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGCTTTTGTGTTT 1178
Qy 841 ggaagaaatccagagtggtgagctgtaaaccttctaaacagagagaggttaagattatacagc 900
Db 1179 GGAAGAAATCCAGAGTGGTGTGATCTGAACCTTCTTAACAGAGAGAGGTAAAGATTATACAGC 1238
Qy 901 tgcacactgtaacttctcagtgctccataaggaagaaactaaagaaacagacgataccatt 960
Db 1239 TGCACACTGTGTAATCTCTAGTGTCCATAGGGGAAGAACTTAAGAGAACCGATACCAAT 1298
Qy 961 ttctggccaggttctctctgtttaaacgctgtgtggaactgtgcctgttgtctccac 1020
Db 1299 TTCGTGGCCAGGTGCTCTCTCTGTTAAACGCTGTGTGTGGAACTGTGCTGTGTCTCCAC 1358
Qy 1021 aattgcaatgaatgcaatgttcccaagcaaaagttaataaaaaataccacaggtcctt 1080
Db 1359 AATTGCAATGAATGCAATGTGTCCTCAAGCAAAAGTTACTAAATAATACACAGGTCCTT 1418
Qy 1081 cagttgagacaaagacggtgtcagggattgcacaaatcactcaccagcgtgcccctg 1140
Db 1419 CAGTTGAGACCAAGACCGGTGTGAGGGGATGCAAAATCACTCACCAGCTGGCCCTG 1478
Qy 1141 gagcaccatgaggagtgactgtgtgcagagaggagacagaggagatagccgcatcac 1200
Db 1479 GAGCACCATGAGGAGTGTACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGATCAC 1538
Qy 1201 caccagagctcttgcacagagctgtgcagtllygcagtggtgctatttatagagaac 1260
Db 1539 CACAGCAGCTCTTGCACAGAGCTGTGCAGT- ---GCAGTGGCTGATTTCTATTAGAGAAC 1594
Qy 1261 gtatgcttatctccatcttaactcagttgttggcttgcagacaccttcatctcagg 1320
Db 1595 GTATGCGTATATCTCATCTCTAATCTCAGTTGTGTGTTCAGGAGACCTTTCACTCTCAGG 1654
Qy 1321 attacagtgctattctgaaagagagacatcaaacagaaattaggagttgtgcaacagctc 1380
Db 1655 ATTTACAGTGCATCTGAAGAGAGAGACATCAACAGAAATTAGGAGTTGTGCAACAGCTC 1714
Qy 1381 ttttgagagaggcctaaagagacaggaagaggtcttcaactcgtggaaagaaataaa 1440
Db 1715 TTTTGAGAGAGGCGCTAAAGACAGGAGAAAGGTCTTCAATCTGTGAAAGAAATATAA 1774
Qy 1441 tgtgttaataatagatcaccagctagtctcagagttaccatgtacgtattccacatagct 1500
Db 1775 TGTGTATTAATAATAGATCACCAGTAGTCTTCAGAGTTTACCATGTATTCACATAGCT 1834
Qy 1501 ggggtctgtatttcagttcttcgatcagcgttagggtaattgtcagtcacagagaaaaaac 1560
Db 1835 GGGTCTGTATTTCAGTCTCTTCGATACGGGTTAGGGTAATGTCTAGTACAGGAAAAAATC 1894
Qy 1561 tgtgcaagtgagcacctgattccgttgccttgaacttaactaaagcctcattcctcgtggc 1620
Db 1895 TGTGCAAGTGAGCACCTGATTCCGTTGCTTGAATTAACCTTAAGCTCCATGCTCCTGGC 1954
Qy 1621 ctaaaatcgatataaaatctggattttttttttttttttttttttttttttttttttttta 1680
Db 1955 CTAATATCGTATAAAATCTGGA- - - - -TTTTTTTTTTTTTTTTTTTCTCATATTCATATGTA 2013

Qy 1681 accagaaacttctatgtactacaaacctgtttttaaaggaactatgttctgtatgaat 1740
Db 2014 ACCAGAACTTCTATGTACTACAACTGTGTTTTTAAAGGAACATGTGTGCTATGAAT 2073
Qy 1741 taaactgtgtgctgtgatagga 1764
Db 2074 TAAACTGTGTGCTGTGTATGGA 2097
RESULT 3
AX047650 2849 bp DNA PAT 15-DEC-2000
LOCUS Sequence 9 from Patent WO0070050.
DEFINITION AX047650
ACCESSION AX047650
VERSION AX047650.1 GI:11876693
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2849)
AUTHORS Baker, K.P., Chen, J., Ferrara, N., Fong, S., Goddard, A., Gurney, A.L., Hillan, K.J., Kuo, S.S., Tumas, D. and Wood, W.I.
TITLE Compositions and methods for the treatment of immune related diseases
JOURNAL Patent: WO 0070050-A 9 23-NOV-2000;
FEATURES Location/Qualifiers
source 1..2849
/organism="Homo sapiens"
/db_xref="taxon:9606"
2715
unsure /note="unknown base"
BASE COUNT 851 a 528 c 619 g 850 t 1 others
ORIGIN
Query Match 98.2%; Score 1732.4; DB 9; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
Qy 1 attatgtgaaactaccctgcgattctctgtccagagcaggtcggcgcttccacccc 60
Db 132 ATTATGTGAAACTACCTCGGATTCTCTGTCCAGAGCAGGCTGGCGCTTCCACCCC 191
Qy 61 agtgcagcttccctggcggtgtgtaagagagactcggagagtcgcttccaaagtgc 120
Db 192 AGTGCAGCTTCCCTTGGCGGTGGTGAAGAGACTCGGAGTCTGCTTCCAAAGTGCC 251
Qy 121 cgcgtgagtgagctctcaccctcagtcacaaatgagcctcttcgggcttctcctgctg 180
Db 252 CGCCGTGAGTGAGCTCTCACCCTCAGCCCAATGAGCTCTTCGGGCTTCTCTGCTG 311
Qy 181 acatctgcctggccgcccagagacaggggactcagcggaatccaaacctgagtagtaa 240
Db 312 ACATCTGCCCTGCGCCGCGCAGAGACAGAGGAGCTCAGGCGGAATCCAACCTGAGTAGTAAA 371
Qy 241 ttccagttttccagcaacaaggaacagagcggagtcacagatcctcagatgagagaatt 300
Db 372 TTCCAGTTTTTCCAGCAACAAGGAACAGACGGAGTACAAGATCCTCAGCATGAGAAAT 431
Qy 301 attactgtctactaataagagattctcagccccaggttttctctcactatcttcaaga 360
Db 432 ATTACTGTCTTACTAATGAAGTATTCAGCCCCAAGGTTTCTCTCATCTTATCCAAGA 491
Qy 361 aatacggcttctggtatggagatttagcagtagagagaaatgatggatacaacttacg 420
Db 492 AATACGGTCTTGGTATGGAGATTAGTAGCAGTAGAGGAAAAATGATGATACAACTACG 551
Qy 421 ttgatgaagatttgggttgaagaccagagaagatgcacatatgcgaagtatgatttga 480
Db 552 TTTGATGAAGATTTGGGCTTGAAGACCCAGAGATGACATATGCAAGTATGATGTTTGA 611


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="embryo"
/tissue_type="brain"
327..1364
/gene="hscdgp"
CDS
327..1364
/gene="hscdgp"
/codon_start=1
/product="spinal cord-derived growth factor"
/protein_id="BAB03266.1"
/db_xref="GI:9392294"
/translation="MSLFGLLITSLAQGRQGTQASNSKQFSSNKBNQVQDP
QHERITTVESGSIHSPFPHYPRNTVLVRLVAVENWMIQTFDFERFLEDPEDD
ICKYDFVEVEPSDGTILGRWCSGTVPKGQISKGNQIRIRFVSDVEYFPSPGFIHY
NIMVQFTVAVSPVLPSPSALPDLLNNAITAFSTLEDLIRFLEPWRWOLDEDLYRP
TWOLLGKAFVGRKSRVDLNLLEEVRLYSCIPRNFVSISRELEKTDRTDFWPGCLL
VRKGGNCACCLHNCNECQVPSKYTKYHEVQLRPKTVGRGLKSLTDVALEHHEE
CDCVCRGSGG"
BASE COUNT      501 a   412 c   424 g   477 t      3 others
ORIGIN

Query Match      89.2%; Score 1574; DB 85; Length 1817;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1629; Conservative 0; Mismatches 10; Indels 9; Gaps 4;

QY 1 attatgtgaaactaccctcgattctctgctgcacagcagcctcgcgcttccacccc 60
Db 174 ATTAATGTGGAAACATACCCCTCGGATCTCTGCTGCCAGAGCAGGCTCGCGCTTCCACCCC 233

QY 61 agtcagccttccctcgcggtggtgaaagactcggagtcgctgcttccaaagtgcc 120
Db 234 AGTCAGCCTTCCTCGCGTGTGAAAGAGACTCGGGAGTCGCTGCTTCCAAAGTGCC 293

QY 121 cgcgctgagtgagctctcaccctcagtcagcgaatagcctctctgctctctctg 180
Db 294 CGCGCTGAGTGAGCTCTACCCCTCAGTCCAGCCAAATGAGCCTCTTCGGGCTTCTCGTGTG 353

QY 181 acatctccctgcccgcagagacaggggaactcaggcgggaatccaaactgagtagtaa 240
Db 354 ACATCTGCCCTGGCGCCAGAGACAGGGAGTCTCAGCGGAATCCAACTGAGTAGTAAA 413

QY 241 ttccagtttccagcaacaggaacgaacgcgaggtacagatcctcagcatgagagaatt 300
Db 414 TTCCAGTTTCCAGCAACAGGAACAGACGAGGTACAGATCCTCAGCATGAGAGATTT 473

QY 301 attactgtctactaatggaagtattccacagcccaaggtttcctcactatccaaga 360
Db 474 ATTACTGTCTACTAATGGAAGTATTACACAGCCCAAGGTTTCCTCATACTATATCCAAGA 533

QY 361 aatacgtcttgtagagattagtagcagtagaggaataatgatgatacaacttacg 420
Db 534 AATACGGTCTGGTATGAGAGATTAGTAGCAGTAGAGGAAATGATGGATACAACTTACG 593

QY 421 ttgatgaagatttggcttgaagaccaggaagatcacatgacatgacatgattgttga 480
Db 594 TTGTGATGAAGATTGGGCTTGAGACCCAGCAAGATGACATGACATGCAAGTATGATTTGTA 653

QY 481 gaagttgaggaaccagtgatggaactatataggcgctggtggtggttctggtactgta 540
Db 654 GAAGTTGAGGAACCCAGTGATGGAACATATTAGGGCGCTGCTGGTGTCTGGTACTGTA 713

QY 541 ccaggaacacagatttctaaaggaaatcaaataggataagattgttatctgatgaatat 600
Db 714 CCAGGAACACAGATTCTTAAAGGAAATCAAAATTAGGATAAGATTGTATCTGATCAATAT 773

QY 601 ttctctctaaaccaggttctgcatccactacacaaattgtcatgccaaattcacagaa 660
Db 774 TTTCCTTCTGAACACAGGGTCTGTGATCCACTACAAACATTTGTCATGCCACAATTCACAGAA 833

QY 661 gctgtgagctcttcagtgctaccctctcagcttctgacactggacctgcttaataatgct 720

```

```

Db 834 GCTGTGAGTCCCTTCAGTGTGTACCCCTTCAGCTTTTGCCACTTGACCTGCTTAATATGCT 893
QY 721 ataaactgctttagtaccttgaagaccttattcgatatcttgaaccagagagatggcag 780
Db 894 ATAACTGCTTTAGTACCTTGGAAAGACCTTATTCGATATCTTGAACACAGAGATGGCAG 953
QY 781 ttggacttagaagatctatataggccaacttggcaacttcttggcaagccttttgtttt 840
Db 954 TTGGACTTTAGAAGATCTATATAGGCCCACTTGCAACTTCTTGCGAAGCTTTTGTGTTTT 1013
QY 841 ggaagaaatccagagtggtgatctgaaccttctaacagagagaggaagataatacagc 900
Db 1014 GGAAGAAATCCAGAGTGTGTGATCTGAACCTTCTAACAGAGAGGTAAGATTATACAGC 1073
QY 901 tgcacacctcgttaactctcagtgctccataagggagaactaaagagaaccatcaccatt 960
Db 1074 TGCACACCTCGTAACTTCTCAGTGTCCATPAGGGAAGAACTTAAGAGAACCCGATACCAT 1133
QY 961 ttctggccagggttgctctctcgtgttaaacgctgtggtggaaactgtgcttgttctccac 1020
Db 1134 TTCTGCCAGGTTGCTCTCTGTTAAACGCTGTGTGGTGGAACTGTGCTGTTGTTCTCCAC 1193
QY 1021 aattgcaatgaatgtaattgtgtcccaagcaagttaactataaaataccacagagtcctt 1080
Db 1194 AATTGCAATGAATGCAATGTGTCTCCAGCAAGTTTACTTAAAAAATACCACGAGTCTCT 1253
QY 1081 cagttgagaccacaaacacggtgtcaggggattgcacaaactcactcacgacgtgcccgtg 1140
Db 1254 CAGTTGAGACCAAAAGACCGGTGTACGGGATTGACAAATCACTACCACGCTGGCCCTG 1313
QY 1141 gagcaccatgagagtgtagactgtgtgcagagggagcacagggagtagcgcgcatcac 1200
Db 1314 GAGCACCATTGAGGAGTGTACTGTGTGTCAGAGGGGACAGGAGGATAGCGCATCAC 1373
QY 1201 caccagcagctcttgcacagagctgtgcagtllylncagtggtgattctattagagaac 1260
Db 1374 CACCAGCAGCTTTGCCCAAAGCTGTGCAGT---GCAGTGGCTGATTCTATTAAAGAAC 1429
QY 1261 gtatcggttatctcaatccttaactcagttgttcttcaaggaccttctcattcagg 1320
Db 1430 GTATCGGTTATCTCCATCCTTAATCTAGTTGTTTGTCTTCAAGGACCTTTTCATCTCAGG 1489
QY 1321 attcagtgctattctgaaagaggagacatcaaacagaattagagattgtgcaacagctc 1380
Db 1490 ATTTACAGTGCATCTTGAAAGAGGAGACATCAACAGAGATTAGGAGTTGTGCAACAGCTC 1549
QY 1381 ttttgagaggagccttaagagcagagagaaagggttctcaatcgttggaagaaaaattaaa 1440
Db 1550 TTTTGAGAGGAGGCGCTTAAAGGACAGAGAGAAAGGTTCTCAATCGTGGAAAGAAAATAAA 1609
QY 1441 tttgtattaaatagatcaacagctagtttcaagatttaaccatg-tacgtattccaactagc 1499
Db 1610 TGTGTATTAATAATAGATCACCCAGCTAGTTTTCAGAGTTTACCATGTTATTCCTACTAGC 1669
QY 1500 tgggtctgtattcagttcttctcagatcagcgttagggttaat-gtcagttacaggaaaaa 1558
Db 1670 TGGGTTCTGTATTTCAGTTCCTTCGATACAGCTTAGGGTAATAGTCAGTACAGGCAAAA 1729
QY 1559 actgtg---caagtgagcacctgattccgcttgccttgccttaactcctaaagctcccatgtcc 1615
Db 1730 ATACTGTGCAGAGTATCACCCTGATTCCTGCTTGCCTTGCCTTAACTCTAAAGCTCCATGTC 1789
QY 1616 tgggcctaaaaatcgataaaaaatcctggat 1643
Db 1790 TGGGCCTAAAAATCGTATAAAATCTGGAT 1817
```

```

RESULT 6
AX027935 LOCUS AX027935 1035 bp DNA PAT 16-SEP-2000
DEFINITION Sequence 3 from Patent WO0037641.
ACCESSION AX027935
VERSION AX027935.1 GI:10188752
```

KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS Gordon,R.D., Dijkman,J.J., Sprengel,J.J., Yon,J.R., Xu,J., Gosiowska,A. and Dhanaraj,S.N.
TITLE Vascular endothelial growth factor-x
JOURNAL Patent: WO 0037641-A 3 29-JUN-2000;
GORDON ROBERT DOUGLAS (BE) ; DIJKMANS JOSIENA JOHANNA HUBER (BE) ; JANSSEN PHARMACEUTICA NV (BE) ; SPRENGEL JORG JURGEN (BE) ; YON JEFFREY ROLAND (BE) ; XU JEAN (US) ; GOSIEWSKA ANNA (US) ; DHANARAJ SRIDEVI NAIDU (US)
FEATURES Location/Qualifiers
source 1..1035
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 301 a 213 c 251 g 270 t
ORIGIN
Query Match 58.7%; Score 1035; DB 9; Length 1035;
Best Local Similarity 100.0%; Pred. No. 2.9e-266;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 154 atgagcctctcggtctcctgctgacatctgcccctggccgagacagagggact 213
Db 1 ATGAGCCCTCTCGGGCTTCTCTGCTGACATCTGCCCTGGCCGCGCAGACAGGGGACT 60
Qy 214 caggcgggaatccaacctgagtagtaaatccagtttccagcaacaagaacagaaacgga 273
Db 61 CAGCGGGAATCCAACCTGAGTAGTAATTCAGATTTCAGCAACAAGCAAGACGAGCA 120
Qy 274 gtacaagatcctcagcatgagagaattattactgtgtctactaattggaagtattccagc 333
Db 121 GTACAAGATCTCTCAGCATGAGAGAATTATTACTGTCTCTACTAATGAAGATTATTCACAG 180
Qy 334 ccaaggttctctactattatccaagaatacggctctggtatgagattagtagcagta 393
Db 181 CCAAGTTTCTCTACTATTATCCAGAAATACGGTCTTGATGAGAGATTAGTAGCAGTA 240
Qy 394 gaggaaaaatgatgatacaactacgtttgatgaagaatttgggcttgaagaccagaa 453
Db 241 GAGGAAAATGATGATACAACTACGTTTGATGAAGATTGGGCTTCAAGACCCAGAA 300
Qy 454 gatgacatatgcaagtatgatttttagaagttgaggaacccagtgatgaaactatata 513
Db 301 GATGACATATGCAAGTATGATTTGTAGAAGTTGAGGAACCCAGTGATGGAACATATTA 360
Qy 514 gggcgtggtgtggttctggtactgtaccagaaacagatttctaaaggaaatcaatt 573
Db 361 GGGCGCTGGTGTGGTACTGTACAGAAACACAGATTCTTAAGGAAATCAAAAT 420
Qy 574 aggaataagattgtatcgtatgaattatttctctgaaccagggttctgcactccactac 633
Db 421 AGGATAAGATTGTATCTGATGAATATTTTCTCTGTAACCAAGGTTCTGATCCACTAC 480
Qy 634 aacattgctgcacaattcacagaagctgtgagtccttcagtgctaccccccttcagct 693
Db 481 AACATTGTCATGCCACAATTCACAGAAGCTGTGAGTCTCTCAGTGTGATACCCCTTCAGCT 540
Qy 694 ttgcactggacctcttaataatcgtataactgcttttagtaccttgaagacccttatt 753
Db 541 TTGCACATGGACCTCTTAATTAATGCTATTAATGCTTTTAGTACCTTTGGAAGACCTTATT 600
Qy 754 cgatatcttgaaaccagagatggcagttggacttagaagatcttatatagggcaacttgg 813
Db 601 CGATATCTTGAACACAGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGG 660
Qy 814 caactcttgcaagcctttttgttttttggaaagaaatccagagtggtgagtagtagcactt 873
Db 661 CAACTTCTTGCAAGGCTTTTGTGTTTGGAGAAAATCCAGAGTGTGTGGATCTGAACCTT 720

Qy 874 ctacacagagagaggaattatacagctgcacacacctcgtaactctcagtggttccataag 933
Db 721 CTAACAGAGAGGTAAAGATTATACAGCTGCACACCTCGTAACCTCTCAGTGTCCATAAG 780
Qy 934 gagaactaaagagaaccgatacattttctggcaggttgctctcctgggttaaacgtgt 993
Db 781 GAAGAATTAAGAGAACCGATACATTTCTGGCCAGGTTGTCTCTCGTTAAACGCTGT 840
Qy 994 ggtggaactgtgcctgttctcccaaatgcaatgaatgaatgtcaatgtgcccaagcaa 1053
Db 841 GGTGGGAATCTGCTCTGTCTCCCAATTGCAATGCAATGAATGTCAATGTGTCCTCCCAAGCAA 900
Qy 1054 gttactaaaaataccacagaggtccttcagttgagacacaaagaccggtgtcagggattg 1113
Db 901 GTTACTAAAAAATACCACAGAGGTCCTTCAGTTGAGACCAAGACCGGTGTGAGGGATTG 960
Qy 1114 cacaatcactcaccgacgtggccctggagcaccatgagagtgtagtgactgtgtgcaga 1173
Db 961 CACAAATCACTACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGA 1020
Qy 1174 gggagcacagagga 1188
Db 1021 GGGAGCACAGGAGGA 1035
RESULT 7
AX028032 AX028032 1035 bp DNA PRI 16-SEP-2000
LOCUS Sequence 3 from Patent WO0037641.
DEFINITION AX028032
ACCESSION AX028032
VERSION AX028032.1 GI:10188844
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS Gordon,R.D., Dijkman,J.J., Sprengel,J.J., Yon,J.R., Xu,J., Gosiowska,A. and Dhanaraj,S.N.
TITLE Vascular endothelial growth factor-x
JOURNAL Patent: WO 0037641-A 29-JUN-2000;
GORDON ROBERT DOUGLAS (BE) ; DIJKMANS JOSIENA JOHANNA HUBER (BE) ; JANSSEN PHARMACEUTICA NV (BE) ; SPRENGEL JORG JURGEN (BE) ; YON JEFFREY ROLAND (BE) ; XU JEAN (US) ; GOSIEWSKA ANNA (US) ; DHANARAJ SRIDEVI NAIDU (US)
FEATURES Location/Qualifiers
source 1..1035
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 301 a 213 c 251 g 270 t
ORIGIN
Query Match 58.7%; Score 1035; DB 91; Length 1035;
Best Local Similarity 100.0%; Pred. No. 2.9e-266;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 154 atgagcctctcggttctcctgctgacatctgcccctggccgagacagagggact 213
Db 1 ATGAGCCTCTCGGGCTTCTCTGCTGACATCTGCCCTGGCCGCGCAGACAGGGGACT 60
Qy 214 caggcgggaatccaacctgagtagtaaatccagtttccagcaacaagaacagaaacgga 273
Db 61 CAGCGGGAATCCAACCTGAGTAGTAATTCAGATTTCAGCAACAAGCAAGACGAGCA 120
Qy 274 gtacaagatcctcagcatgagagaattattactgtgtctactaattggaagtattccagc 333
Db 121 GTACAAGATCTCTCAGCATGAGAGAATTATTACTGTCTCTACTAATGAAGATTATTCACAG 180
Qy 334 ccaaggttctcactactattatccaagaatacggctctggtatgagattagtagcagta 393
Db 181 CCAAGTTTCTCTACTATTATCCAGAAATACGGTCTTGATGAGAGATTAGTAGCAGTA 240

QY 394 gaggaataatgatacacaacttaactgcttgatgaagattggcttgagaccagaa 453
Db 241 GAGGAAATCTATGGATACAACTTACGTTGATGAAAGATTGGCTTGAAGACCCAGAA 300
QY 454 gatgacatatgaagatgattttgttagaagttgagaaacccagtgagaaactatata 513
Db 301 GATGACATATGCAAGTATGATTTTGTAGAAGTTGAGAAACCCAGTGATGGAACATATATTA 360
QY 514 ggcgcctgggtgctgctgactgtaccaggaacacagatttttaaggaatacaaat 573
Db 361 GGGCGCTGGTGTGCTGTGCTACTGTACCAAGAAACAGATTTCTAAGGAAATCAAAAT 420
QY 574 aggataagatttgatgatgaatatttccctctgaacacaggttctgtcatccactac 633
Db 421 AGGATAAGATTGTATCTGATGAATATTTCTTCTGAACAGGTTCTGTGCATCCATAC 480
QY 634 aacattgtcatgccacaattcaagaagctgtgagctcctcaagtgtaccaccttcagct 693
Db 481 AACATTGTCTATGCCACAAATTCACAGAAGCTGTGAGTCTCTCAGTGTACCCCTTCAGCT 540
QY 694 ttccactgacactgcttaataatgctataactgctttagtaccttggagacattatt 753
Db 541 TTGCCACTGGACCTGCTTAATATGCTTAACTGCTTGTAGTCTTGTAGAGACCTTATT 600
QY 754 cgatatcttgaaccagagatggcagttgagcttagaagatctatataggccaacttgg 813
Db 601 CGATATCTTGAACACAGAGATGCGACTTGGACTTTAGAAGATCTATATAGGCCAACTTGG 660
QY 814 caacttctggcaagcttttggtttttgggaagaaatccagagtggtgagctgaacctt 873
Db 661 CAACCTTCTTGGCAAGGCTTTTGTGTTTGGGAAGAAATCCAGAGTGTGATCTGAACCTT 720
QY 874 ctacagagagatgaattatcacagctcacacctgctgaactctcaagtccatgaagg 933
Db 721 CTACAGAGAGGTAAGATTATACAGCTGCACACCTCGTAACCTTCTCAGTGTCCATAAGG 780
QY 934 gaagaactaaagagaaccgatacattttctggcaggtgtctcctgtttaacgcgtgt 993
Db 781 GAAGAATAAAGAGAACCGATACATTTTCTGGCCAGGTTGTCCTGTTAAACGCTGT 840
QY 994 ggtgggaactgtgctgtgtctccacaaatgcaatgaatgtcaatgtgtcccaagcaaa 1053
Db 841 GGTGGGAACCTGTGCTGTCTCTCCACAAATGCAATGCAATGCAATGTGTCCCAAGCAA 900
QY 1054 gttactaaaaataaccagaggtccttcagttgagacaaacagcgtgtcagggattg 1113
Db 901 GTTACTAAAAAATACCACGAGGCTTTCAGTTGAGACAAAGACCGGTGTCCAGGGGATTG 960
QY 1114 cacaatacctcaccgacgtggcctgagacaccatgagagtgtagtgtgtgtgcaga 1173
Db 961 CACAAATCACTACCGACGCTGGCCCTGAGACACCATGAGGAGTGTGACTGTGTGTCAGA 1020
QY 1174 gggagcacaggagga 1188
Db 1021 GGGAGCACAGGAGGA 1035

RESULT 8
AX044520
LOCUS AX044520 3571 bp DNA PAT 24-NOV-2000
DEFINITION Sequence 34 from Patent WO0066736.
ACCESSION AX044520
VERSION AX044520.1 GI:11343375
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3571)
AUTHORS Gilbert,T., Hart,C.E., Sheppard,P.O. and Gilbertson,D.G.
TITLE Growth factor homolog zveg4
JOURNAL Patent: WO 0066736-A 34 09-NOV-2000;

ZymoGenetics, Inc. (US)
Location/Qualifiers
1. 3571
/organism="Mus musculus"
/db_xref="taxon:10090"
1049. .2086
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC17171.1"
/translation="MLLGLLLLSALAGORTGTRAESNLSSKLLQLSSDKQNGVQDD
RHERVVISNGSHSPKFPHTYPRNMVLYAVDENVRIQJTFDRFLEDPEDD
ICKYDFVEEPPSGVLRGWCSTGKQTSKGNHIRIRFVSEYSPGFCIH
SIIMPOVETTSPLVPSLSLLDNNAVTAFLEELIRYLPDRQVLDLSLYK
TWOLLKRAFLYKKSKVNLNLLKEEVLYSCTPRNFVSIRLELKRDTDTLPWGLL
VKRGNCACCLHNCQCQVPRKVTYKHYEVQLRPKTYKVKGLHKSITDVALEHHEE
CDCVCRGNAGG"
BASE COUNT 876 a 935 c 875 g 885 t
ORIGIN
Query Match 50.1%; Score 883.2; DB 9; Length 3571;
Best Local Similarity 83.7%; Pred. No. 1.5e-225;
Matches 1014; Conservative 0; Mismatches 193; Indels 5; Gaps 1;
QY 1 attatgtgaaactaccctgcgattctctgtgcagagcagctgcgcgttcacccc 60
Db 901 ATTATGTGAAACTACCTGCGATTCTCTGCTGCCAGAGCCGCCGCTTCCACCGC 960
QY 61 agtcagacctccccggcgtggtgaaagagactcgggagtcgctgtcttcaaaagtgc 120
Db 961 AGCGCAGCGCTTCCCGCGCTGGGCTGA-----GCCTTGGAGTCTGCTGCCAGTGC 1015
QY 121 cgcctgagtgagctctcaacccagtcagcgaataatgagcctcttctgggtctctctgtg 180
Db 1016 CGCCCGAGTGTAGCCTCTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
QY 181 acatctgcctgtgcgcgcagagagcagggactcagcggaatccaaactcagtgagtaaa 240
Db 1076 ACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
QY 241 ttccagtttccagcaacaaagaaacagacgaggtacagatcctcagcatgagagaatt 300
Db 1136 TTGCAGCTCTCCAGCGACAAAGAACAGACGAGTGAAGATCCCGGCTGAGAGAGTT 1195
QY 301 attactgtctactaataatggaagtattcacgcccaagtttccctctacttactatcaaga 360
Db 1196 GTCACTATATCTGTGTAATGGGAGCATCCACGCCGGAAGTTTCTCTCATACATACCAAGA 1255
QY 361 aatcaggtctgtgtgagatttagcagtagaggaataatgtagatacaacttaccg 420
Db 1256 AATATGGTGTGCTGGTGGAGATTAGTTGAGTAGATGAAATGTGCGGATCCAGCTGACA 1315
QY 421 ttgtatgaagatttggccttgaagaccagaaagatgacatatgcaagtattgtttgta 480
Db 1316 TTTGATGAGAGATTGGCTGGAAGATCCAGAAAGACGATATATGCAAGTATGATTTGTA 1375
QY 481 gaagttaggaacccagtgatggaactatttagggcgtgtgtgtgtgtgtgtgtgtgtgt 540
Db 1376 GAAAGTTGAGGACCCAGTGATGGAAGTGTGTTAGGACGCTGTGTGTGTGTGTGTGTGTGT 1435
QY 541 ccaggaaacagattcttaaggaaatcaaattagataagattgtatctgtatgaatat 600
Db 1436 CCAGGAAAGCAGACTTCTAAAGAAATCATATCAGGATAAGATTGTGATCTGATGAGTAT 1495
QY 601 ttctctctgaacaggttctgtccactacaacattgcatcccaaatccacagaa 660
Db 1496 TTTCCATCTGAACCCGGGATCTGCAATCCACTACATGATATATCATGCCACAGTACAGAA 1555
QY 661 gctgtgactcttcagtgctaccctctcagcttggcagctgacctgaccttaataatgct 720
Db 1556 ACCAGAGTCTCTCGGTGTGTCCTTCACTTTGTCTATGGACCTGCTGCACAACTGCT 1615

```
QY 721 ataactgccttagtaccttggaagacccttattcgatatctcttgaaacacagagagatggcag 780
Db 1616 GTGACTGCTTACAGTACCTTGAAGAGAGCTGATTCGGTACTACAGCAGCATGCATGCAG 1675
QY 781 ttggacttagaagatctatataggcccaactgtgcaactctctggcaagcctttgtttttt 840
Db 1676 GTGGACTTGGACAGCCTCTACAGCCACATCGCAGCTTTTGGCAAGGCTTTCTCTGTAT 1735
QY 841 ggaagaaatccagagtggtgagctgaaccttcttaacagagagaggtgaagattatcacgc 900
Db 1736 GGGAAAAAAGCAAGTGGTGAATCTGAATCTCTCAAGAGAGAGGTAAACCTCTACAGC 1795
QY 901 tgcacacctgttaactctcaagtgtccataaggggaagaaactaaagaaacccgataccatt 960
Db 1796 TGCACACCCCGGAACCTTCTCAGTGTCCATACGGAAGAGCTAAAGAGGACAGATACCATA 1855
QY 961 tcttgccaggtgtctctgtttaaagcgtgtgtggaactgtgcctgtgtcttccac 1020
Db 1856 TTCTGGCCAGGTTGCTCTCTGTCAAGCGCTGTGGAGGAATTTGTGCCCTGTCTCTCCAT 1915
QY 1021 aattgcaatgaatgtcaatgtgtcccaagcaaaagtgtactaaaaataaccacagaggtcctt 1080
Db 1916 AATTGCAATGAATGTCAAGTGTCCACGCTAAAGTTACAAAAAAGTACCATGAGGTCTT 1975
QY 1081 cagttgagacaaagacccggtgtcaggggattgcacaaatcactaccacgacgtggccctg 1140
Db 1976 CAGTTGAGACCAAAACCTGGAGCTCAAGGGATTGCATAGTCACTCACTGATGTGGCTCTG 2035
QY 1141 gagcacctagagagtgactgtgtgcagagggagcacagagagatagccgacatcac 1200
Db 2036 GAACACACGAGGAATGTGCTGTGTGTAGAGGAACGCGAGGGTAACTGCGACCT 2095
QY 1201 caccagcagctc 1212
Db 2096 TCGTAGCAGCAC 2107

RESULT 9
AF117608 2692 bp mRNA ROD 02-JAN-2000
DEFINITION Mus musculus fallotein mRNA, complete cds.
ACCESSION AF117608
VERSION AF117608.1 GI:6652867
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Tsai,Y.-J., Lee,R.K.-K., Chen,Y.-H., Lin,S.-P. and Cheng,W.T.-K.
TITLE cDNA cloning of fallotein from mouse ovary
JOURNAL Unpublished
REFERENCE 2
AUTHORS Tsai,Y.-J., Lee,R.K.-K., Chen,Y.-H., Lin,S.-P. and Cheng,W.T.-K.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1999) Medical Research, Mackay Memorial Hospital, 45 Min Sheng Road, Tamshui, Taipei 25115, Taiwan
FEATURES
SOURCE 1..2692
/organism="Mus musculus"
/db_xref="taxon:10090"
/sex="female"
/tissue_type="ovary"
/dev_stage="adult"
198..1235
/note="putative secretory protein"
/codon_start=1
/evidence-not_experimental
/product="fallotein"
/protein_id="AAF22516.1"
/db_xref="GI:6652868"
/translacion="MLLGLLLLTSLAGQRTGFAESNLSSKILQLSDKRGNGVQDP
RHERVVTISNGSIHSPKFPHTYPRNMLVWRLVAVDENVRIQLTFDRFGLDEPDD
```

```
polyA_signal 754 a 576 c 617 g 745 t
BASE COUNT 2662..2667
ORIGIN
ICKYDFVEVEEESDGSVLGRWCGSVLPGKQTSKGNHIRFVSDYFFSEPGFCIHY
SIIMQVETTSPLSPSSLDLNNAVTAFSTLEELIRYLEPWRQWDDLSLYKP
TMOLLKAFLYGKSKVNLNLLKEEVLYSCTPRNFVSIRLELRKTDIFWPGLL
VKRCGNCACCLHNCNECCQPKVTKKYHEVLQLRPKTKVGLHSLTDVALEHHEE
CQCVCRNAGG"
2662..2667
Query Match 50.0%; Score 881.6; DB 94; Length 2692;
Best Local Similarity 83.6%; Pred. No. 4e-225;
Matches 1013; Conservative 0; Mismatches 194; Indels 5; Gaps 1;
QY 1 attatgtggaactaccctgcattctctgctgcagagcagcgtcgcgtcttcacccc 60
Db 50 ATTATGTGGAACCTACCTCGGATTCTCTGCTGCCAGAGCGGCCAGCGCTTCCACCGC 109
QY 61 agtcagcgtctccctgcgtgcgtggtgaaagagactcgggagtcgctctccaaagtcc 120
Db 110 AGCGCAGCCTTCCCGCGCTGGGCTGA- - - - -GCCTTGGAGTCTGCTTCCCGAGTGC 164
QY 121 cgcgtgagtgagctctcaccacagtcagccaaatgagcctcttcgggctctctctg 180
Db 165 CGCGCGGAGTGAGCCTCGCCCGCAGTCAGCCAAATGCTCTCTCGGCCCTCTCTCTG 224
QY 181 acatctgcctcggcggccagagacagggactcagcgggaatcccaacctgagtagtaa 240
Db 225 ACATCTGCCCTGCGCGCCAAAGACGGGACTCGGGCTGAGTCCCACTTGACGACGAG 284
QY 241 ttccagttttccagcaacaaaggaaacagcggagtagacaagatcctcagcagagaaat 300
Db 285 TTGCAGCTCTCCAGCGACAAAGGAACAGACGGAGTGCAGATCCCGGCATGAGAGATT 344
QY 301 attactgtctactaatgaagattatccacagcccaagggttctcctaactatccaaga 360
Db 345 GTCATATATCTGTTAATGGAGCATCCACAGCCCGGAAGTTTCTCTACATGATACCCAGA 404
QY 361 aatacagctgtgtatgagattagtagcagtagagagaaatgtatgatacaactcacg 420
Db 405 AATATGGTGTGTGTGGAGATTAGTTGAGTAGTAGTAAATGTGCGGATCCAGCTGACA 464
QY 421 ttgatgaagattttgggttgaaagaccagagatgacatagcaagtagatattttgta 480
Db 465 TTTGATGAGAGATTGGGCTGGAAGATCCAGAAGACGATATATCAAGTAGTATTTTGTGA 524
QY 481 gaagtggaaacccagtgatgaactatattaggcgctgggtggtctgtactgta 540
Db 525 GAAGTTGAGGAGCCCACTGATGGAAGCGTTTATAGGACGCTGGTGTCTGGGACTGTG 584
QY 541 ccaggaaacagatttctaaagaaatcaaattagataagattgtctatctgatgaat 600
Db 585 CCAGGAAAGCAGACTTCTAAAGGAATCATATCAGGATAAGATTGTATCTGTATGAGTAT 644
QY 601 ttctcttgaaacaggggttctgcatccatcaacattgtcgcacaaattccagaa 660
Db 645 TTTCCATCTGAACCCGAGTTCTGCATCCACTACAGTATTATCATGCCACAAGTCACAGAA 704
QY 661 gctgtgagctctcagtgctaccccttccagcttggccactggacctgcttaataatgct 720
Db 705 ACCACGAGTCTTTCGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 764
QY 721 ataactgccttagtagcttggagacaccttctcgatatctcttgaaacacagagatggcag 780
Db 765 GTGACTGCTTACAGTACCTTGAAGAGAGCTGATTCGGTACCTAGAGCCAGATCATGTCAG 824
QY 781 ttgacttagaagatcttatagggcaacttggcaactctcttggaaggctttgtttttt 840
Db 825 GTGACTTGGACAGCCTCTACAAAGCAACATGCGAGCTTTTGGCAAGGCTTTCTCTGTAT 884
QY 841 ggaagaaaaatccagagtggtgagatctgaaacctctcaacagagagaggtgaagatatacagc 900
Db 885 GGGAAAAAAGCAAGAGTGGTGAATCTCTCTCTCAAGGAAGAGGTTAAACCTCTACAGC 944
```

QY 901 tgcacacctegtaactctcagtggtccatcaagggaagaactaaagaaacgagacacatt 960
Db 945 TGCACACCCCGGAACCTCTCAGTGTCCATACGGGAAGAGCTAAAGAGGACAGATACCATA 1004
QY 961 ttctggccaggttgctcctcgtgtaaacgctgtggtgggaactgtgctgtgtgtccac 1020
Db 1005 TTTGCGCCAGGTGTCTCCGGTCAACGCTGTGGAGAAATTTGCTGTGTGTCTCCAT 1064
QY 1021 aattgaatgaatgaatgtgtcccaagcaagttactaaaaataaccacagagtcctt 1080
Db 1065 AATTGAATGAATGTAGTGTGTCCACGTTAAAGTTACAAAAAGTACCATGAGTCCCT 1124
QY 1081 cagttgagacccaagacggtgtcaggggattgcacaaatacaactcacgagctggccctg 1140
Db 1125 CAGTTGAGACCAAAAACCTGGAGTCAAGGGATTGCATAGTCACTGATGTGGCTCTG 1184
QY 1141 gacacacatgagagtgctgactgtgtgcagagggagcagagagatagccacatcac 1200
Db 1185 GAACACACAGGAATGTGACTGTGTGTGTAGAGAAACGAGGAGGTAACCTGCAGCCT 1244
QY 1201 caccagcagctc 1212
Db 1245 TCCTAGCAGCAC 1256

RESULT 10
AB033830
LOCUS
DEFINITION
AB033830 1116 bp mRNA ROD 10-FEB-2001
Rattus norvegicus rScdglf mRNA for spinal cord-derived growth factor, complete cds.
ACCESSION
AB033830
VERSION
AB033830.1 GI:11994799
KEYWORDS
spinal cord-derived growth factor; SCDGF.
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (sites)
Hamada,T., Ui-Tel,K., Imaki,J. and Miyata,Y.
Molecular Cloning of SCDGF-B, a Novel Growth Factor Homologous to SCDGF/PDGF-C/fallotin
Biochem. Biophys. Res. Commun. 280 (3), 733-737 (2001)
11162582
2 (bases 1 to 1116)
Hamada,T., Ui-Tel,K. and Miyata,Y.
Direct Submission
Submitted (25-OCT-1999) Tsuyoshi Hamada, Nippon Medical School, Department of Pharmacology; 1-1-5, Sendagi, Bunkyo-ku, Tokyo, 113-8602, Japan (E-mail:t-hamada@nms.ac.jp, Tel:81-3-3822-2131(ex.5277), Fax:81-3-5814-1684)
Location/Qualifiers
1. .1116
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="Kidney"
/dev_stage="Adult"
1. .1038
/gene="rScdglf"
1. .1038
/gene="rScdglf"
/codon_start=1
/product="spinal cord-derived growth factor"
/protein_id="BAB19569.1"
/db_xref="GI:11994800"
/translation="MLLLLLLSALAGORTGTRAESNLSSKLQLSSDKQEVQDP
RHERVVTISGNSIHSPKPPHTYPRNTVLVRLVAVDENVRIOLTDFERFGLDEDD
LKYDFVEEFPDSGLVRCWCSGTVPKQSTGNHRIKRVSDVEYFSEPGFCIH
STIMPOVETTPSPSALSLDLNNVATSTVEELIRLEPDRWQIDLSLVK
TWPLLGKFLYKSKAVNLNLLKEVKLYSCTPRNFSVSIREEKRTDTIFWPGCLL
VRCGNCACCLHNCNEOCQVPRKVTYKTHEVQLRPKIGVKGLHLKSLTDVALEHHEE

FEATURES
source

gene

CDS

CDCVCRGNTEG*

BASE COUNT 301 a 267 c 286 g 262 t
ORIGIN

Query Match 44.2%; Score 780; DB 94; Length 1116;
Best Local Similarity 83.5%; Pred. No. 5.9e-198;
Matches 885; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 154 atgagcctcttcgagctctcctgctgacatgctccctgagccgagacagagggact 213
Db 1 ATGCTCCCTCTCGGCTCTCCTCTGTGACATCTGCCTGGCGCCGAGAGAACCGGACT 60
QY 214 caggcggaatccaaacctgagtagtaaaatccagtttccagcaacaagaagaaacagacga 273
Db 61 CGGGCTGAGTCCAACTGAGCAGCAAGTTGACGCTCTCCAGCGACAGAGCAGACGA 120
QY 274 gtacaagatcctcagcatgagagaattactgtgtctactaataatggaagtattcacagc 333
Db 121 GTGCAAGATCCCGCATGAGAGAGTTGTCACTATATCTGTTAAATGGAGCATCCACAG 180
QY 334 ccaaggtttcctcatactattccaagaaatacgggtctgtgtagagagattagtagcagta 393
Db 181 CGAAAGTTTCCTCATACCTTACCCAAAGAAATACGGTCTGTGCTGGAGATTAGTGCAGTA 240
QY 394 gaggaataatgtatgatacaacttacgtttgatgaaagatttggccttgaagaccacga 453
Db 241 GATGAAATGTCCGATCCAGCTGACATTTGATGAGAGATTTGGGCTGGAACCCAGAG 300
QY 454 gatgacatatgcaagtgtgctgtgtagaattgtagaagttgaggaacccagtgatggaactat 513
Db 301 GACGATCTATGCAAGTATGATTTGTAGAAGTTGAGGAGCCCGAGTCATGGAAGTGTATTA 360
QY 514 gggcgtggtgtgctgtggtactgtaccagaaacagatttctaaagaaatacaaat 573
Db 361 GGACGCTGTGTGTGTGTGGGACTGTGCCAGGAAAGCAGACTTCTAAAGAAATCATATC 420
QY 574 aggataaattgtatcgtatgaataatttcttctgaacacaggttctgcaccactac 633
Db 421 AGGATAAGATTTGTATCTGATGAGTATTTCCCATCGAACCCGGATTCGATCCACTAC 480
QY 634 aacattgtcatcccaaatccagaagctgtgagtcctcagtgctacccctcagct 693
Db 481 AGTATATCATGCCACAAGTCACAGAACTACGAGTCTCTCAGTGTGGCCCTTCAGCT 540
QY 694 ttgacctggacctgttaataatgtactaactgctttagtaccttggagacaccttatt 753
Db 541 TTGTCTATTGGACCTGCTCAATAATGCTGTGACTGCTTCAGTACTCTGGAAGAGCTGATT 600
QY 754 cgatatcttgaaccagagagatgagcagttgagcttagaagatctatataggccaacttgg 813
Db 601 CGGTTCCCTAGAGCCAGATCGATGGCAGATAGACTTTGGACAGGCTCTACAAAGCAACATGG 660
QY 814 caactctctgcaagcctttttttggaagaaataccagagtggtgatctgaacctt 873
Db 661 CGGCTTTTGGGCAAGCTTTCTCTGTATGGGAAAAAAGCAAGCGGTGAATCTGAACCTT 720
QY 874 ctacaagagaggtgaagattatcacagctgcacacctcgtaacttctcagtgctccataagg 933
Db 721 CTCAAAGAAGAGGTAAACTTTTACAGCTGCACACCCCGGAACCTTCTCTGTGTCCATACGG 780
QY 934 gaagaactaaagaaacccgataccattttctgtgagcaggtgtgtcctggtttaaaccgtgt 993
Db 781 GAAGAGCTAAAGAGGACAGACACCATATTTCTGGCCAGGTTGTCTCCTGCTGAAGCGCTGT 840
QY 994 ggtgggaactgtgctgtgtctccacaatttccaatgaatgtcaatgtgtcccaagcaaa 1053
Db 841 GGAGGAAATGTGCTGTGTGCTCCATTAATTCGAATGAATGTGAGTGTGTCTCCACGTAAG 900
QY 1054 gttactaaaaataaccagaggtctctcagttgagacaaacagccggtgtcaggggatttg 1113
Db 901 GTTACAAAAAGTAGCATTGAGGAGTCTCTCAGTTGACACCAAAAATTTGGAGTCAAGGGATTG 960

VERSION KEYWORDS SOURCE

AC009582.4 GI:7658376
HTG: HTGS_PHASE1; HTGS_DRAFT.
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Birren, B., Linton, L., Nusbaum, C. and Lander, E.
1 (bases 1 to 191314)

AUTHORS

Homo sapiens chromosome 4, clone RP11-154F14

JOURNAL

REFERENCE

2 (bases 1 to 191314)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckerly, R., Ben, J., Brown, A.,
Castle, A., Cerny, J., Collangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArillano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hacos, B., Heaford, A., Horton, L., Howland, J.C., Jones, C., Kann, L.,
Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

TITLE

JOURNAL

COMMENT

Submitted (27-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 28, 2000 this sequence version replaced gi:7381775.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1726

Center clone name: 154_F_14

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 173370 bases at least Q40

Consensus quality: 182196 bases at least Q30

Consensus quality: 185724 bases at least Q20

Insert size: 194000; agarose-fp

Quality coverage: 3.6 in Q20 bases; agarose-fp

Quality coverage: 3.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1155: contig of 1155 bp in length

1156 1255: gap of 100 bp

1256 2550: contig of 1295 bp in length

2551 2650: gap of 100 bp

2651 4185: contig of 1535 bp in length

4186 4285: gap of 100 bp

4286 5873: contig of 1588 bp in length

5874 5973: gap of 100 bp

5974 7991: contig of 2018 bp in length

7992 8091: gap of 100 bp

8092 10978: contig of 2887 bp in length

10979 11078: gap of 100 bp

* 11079 13450: contig of 2372 bp in length
* 13451 13550: gap of 100 bp
* 13551 16422: contig of 2872 bp in length
* 16423 16522: gap of 100 bp
* 16523 18672: contig of 2150 bp in length
* 18673 18772: gap of 100 bp
* 18773 22137: contig of 3365 bp in length
* 22138 22237: gap of 100 bp
* 22238 25136: contig of 2899 bp in length
* 25137 25236: gap of 100 bp
* 25237 28475: contig of 3239 bp in length
* 28476 28575: gap of 100 bp
* 28576 31896: contig of 3321 bp in length
* 31897 31996: gap of 100 bp
* 31997 34238: contig of 2242 bp in length
* 34239 34338: gap of 100 bp
* 34339 39364: contig of 5026 bp in length
* 39365 39464: gap of 100 bp
* 39465 44334: contig of 4870 bp in length
* 44335 44434: gap of 100 bp
* 44435 49563: contig of 5129 bp in length
* 49564 49663: gap of 100 bp
* 49664 57336: contig of 7673 bp in length
* 57337 57436: gap of 100 bp
* 57437 63674: contig of 6238 bp in length
* 63675 63774: gap of 100 bp
* 63775 70889: contig of 7115 bp in length
* 70890 70989: gap of 100 bp
* 70990 78191: contig of 7202 bp in length
* 78192 78291: gap of 100 bp
* 78292 85646: contig of 7355 bp in length
* 85647 85746: gap of 100 bp
* 85747 93984: contig of 8238 bp in length
* 93985 94084: gap of 100 bp
* 94085 103005: contig of 8921 bp in length
* 103006 103105: gap of 100 bp
* 103106 114072: contig of 10967 bp in length
* 114073 114172: gap of 100 bp
* 114173 126541: contig of 12369 bp in length
* 126542 126641: gap of 100 bp
* 126642 140995: contig of 14354 bp in length
* 140996 141095: gap of 100 bp
* 141096 154380: contig of 13285 bp in length
* 154381 154480: gap of 100 bp
* 154481 170279: contig of 15799 bp in length
* 170280 170379: gap of 100 bp
* 170380 191314: contig of 20935 bp in length.

FEATURES

Location/Qualifiers
source
1..191314
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-154F14"
/clone_lib="RPC1-11 Human Male BAC"
misc_feature
1..1155
/note="assembly_fragment"
misc_feature
1256..2550
/note="assembly_fragment"
misc_feature
2651..4185
/note="assembly_fragment"
misc_feature
4286..5873
/note="assembly_fragment"
misc_feature
5974..7991
/note="assembly_fragment"
misc_feature
8092..10978
/note="assembly_fragment"
misc_feature
11079..13450
/note="assembly_fragment"
misc_feature
13551..16422
/note="assembly_fragment"
misc_feature
16523..18672
/note="assembly_fragment"

7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 00:12:33 ; Search time 101.97 Seconds
(without alignments)
10862.199 Million cell updates/sec

Title: US-09-457-066-1
Perfect score: 1764
Sequence: 1 attatgtgaaactaccctg.....cttgtctgctgatagga 1764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*

1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:*

2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:*

3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:*

4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:*

5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:*

6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:*

7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:*

8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:*

9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:*

10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:*

11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:*

12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:*

13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:*

14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:*

15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:*

16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:*

17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:*

18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:*

19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:*

20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:*

21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:*

22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1746	99.0	1760	21	AAC81582 Human zveg3 DNA,
2	1746	99.0	1760	21	AA51498 Human growth facto
3	1744.4	98.9	2668	21	AA51990 Human VEGF-X CDNA.
4	1734	98.3	2839	21	AA47452 Human TANGO 128 CO
5	1732.4	98.2	2779	21	AA52458 CDNA encoding huma
6	1732.4	98.2	2825	20	AA23691 Human VEGF-E DNA.
7	1732.4	98.2	2849	20	AA334296 Human PRO200 nucle
8	1732.4	98.2	2849	21	AA78582 Human PRO200 (UNQ1
9	1732.4	98.2	2849	21	AA88515 Human PRO200 (VEGF
10	1732.4	98.2	2849	21	AA58579 Human PRO200 prote
11	1732.4	98.2	2849	21	AA77621 Human PRO713 CDNA

12	1732.4	98.2	2849	22	AAC90564 Human PRO200 cDNA.
13	1732.4	98.2	2849	22	AAC88962 Human PRO200 codin
14	1732.4	98.2	2849	22	AAC97404 Human angiogenesis
15	1729.2	98.0	2858	21	AZ48599 Bone morphogenic p
16	1718.2	97.4	2776	21	AA471952 Human VEGF-X homol
17	1717.2	97.3	2896	21	AAC64426 Human platelet-der
18	1717.2	97.3	2898	20	AA86352 DNA encoding human
19	1689.4	95.8	2475	21	AA471951 Human RACE generat
20	1662.2	94.2	3087	21	AA18314 Lung cancer associ
21	1616.8	91.7	2794	21	AZ48600 Bone morphogenic p
22	1610.8	91.3	2108	21	AA12523 cDNA encoding plat
23	1497	84.9	1536	21	AA12524 Human VEGF-X DNA i
24	1356	76.9	1473	21	AA471955 Human VEGF-X DNA f
25	1038.4	58.9	1134	21	AA471983 MBP-zVEGF3 fusion
26	998	56.6	1095	21	AA51540 Human VEGF-X DNA f
27	977	55.4	1096	21	AA471985 Human VEGF-X DNA f
28	973	55.2	1134	21	AA471984 Mouse zveg3 DNA,
29	883.2	50.1	3571	21	AAC81583 Murine vascular en
30	883.2	50.1	3571	21	AA51527 cDNA encoding muri
31	878.8	49.8	1474	21	AA12525 Human VEGF-X DNA #
32	871	49.4	1878	21	AA471936 Human VEGF-X DNA #
33	857.4	48.6	1716	21	AA471982 Human VEGF-X DNA #
34	750	42.5	822	21	AA471981 Human VEGF-X DNA #
35	737.2	41.8	1035	21	AA51499 Human growth facto
36	648.2	36.7	1710	21	AA471989 Human VEGF-X DNA e
37	470	26.6	542	21	AA471988 Human VEGF-X DNA e
38	447	25.3	550	21	AA471987 Human VEGF-X CUB-1
39	373.4	21.2	764	21	AA47478 Murine TANGO 128 c
40	350	19.8	500	21	AA471986 Human VEGF-X PDGF-
41	283.4	16.1	289	21	AA47809 Human EST clone DN
42	283.4	16.1	289	21	AA47832 Human EST clone DN
43	279	15.8	279	21	AA47807 Human EST clone DN
44	279	15.8	279	21	AA47828 Human EST clone DN
45	278.6	15.8	284	21	AA47802 Human EST clone DN

ALIGNMENTS

RESULT	1
AAC81582	
ID	AAC81582 standard; DNA; 1760 BP.
XX	
AC	AAC81582;
XX	
DT	09-MAR-2001 (first entry)
XX	
DE	Human zveg3 DNA, SEQ ID NO:32.
XX	
KW	Human; zveg3; zveg4 fusion; growth factor homologue; VEGF/PDGF family;
KW	CUB domain; PDGF-like activity; mitogenic; osteogenic;
KW	neovascularisation; tissue repair; proliferation; differentiation;
KW	liver damage; neurodegenerative; Alzheimer's disease; multiple sclerosis;
KW	periodontal disease; bone fracture; wound healing; ischaemia;
XX	immunomodulation; hepatic; ds.
OS	Homo sapiens.
PN	WO200066736-A1.
XX	
PD	09-NOV-2000.
XX	
PF	03-MAY-2000; 2000WO-US40047.
XX	
PR	03-MAY-1999; 99US-0304216.
PR	10-NOV-1999; 99US-0164463.
PR	04-FEB-2000; 2000US-0180169.
XX	
PA	(ZYMO) ZYMOGENETICS INC.
PI	Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;
XX	
DR	WPI; 2000-687541/67.

DR P-PSDB; AAB48657.
XX Growth factor homologs and the nucleic acids that encode them, useful
PT e.g. for treating liver damage, ischemia, multiple sclerosis and
PT Alzheimer's disease -
XX
PS Claim 25; Page 123-125; 143pp; English.
XX
CC The invention relates to the human growth factor homologue zveg4
CC (AAB48653), and nucleic acids encoding it (AAC81555). Zveg4 is a member
CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
CC growth factor) family. Zveg4 has a growth factor domain (AAB48654)
CC characterised by a PDGF cysteine knot structure, and a CUB domain
CC (AAB48655) which has a beta barrel structure. Zveg4 has PDGF-like
CC activity, having mitogenic activity on fibroblasts, vascular smooth
CC muscle cells and pericytes, and has also been shown to stimulate bone
CC growth. The invention also relates to fusion proteins comprising human
CC zveg4 or fragments thereof, particularly human zveg4/human zveg3
CC fusions; expression constructs and host cells comprising human zveg4
CC nucleic acids; the recombinant expression of human zveg4; an antibody
CC which binds to human zveg4 or a fragment thereof; a method of activating
CC a cell-surface PDGF receptor using a zveg4-derived polypeptide; a
CC method of modulating the proliferation, differentiation, migration or
CC metabolism of bone cells, comprising exposing bone cells to
CC zveg4-derived polypeptides; and a method of detecting a genetic
CC abnormality in the zveg4 gene of a patient. Zveg4 proteins and derived
CC fragments may be used to stimulate tissue development or repair, or
CC cellular differentiation or proliferation. They are particularly used for
CC the treatment or repair of liver damage, and may also be used to
CC modulate neurite growth (e.g., in the treatment of Alzheimer's disease or
CC multiple sclerosis). Due to their osteogenic activity, they may be used
CC in the treatment of periodontal disease and fractures. They may also be
CC used to enhance expansion and mobilisation of haematopoietic stem cells
CC and endothelial precursor stem cells, which may be useful in the
CC treatment of ischaemia, in wound healing, and in the modulation of the
CC immune system. The present sequence represents DNA encoding human
CC zveg3.
XX
SQ Sequence 1760 BP; 494 A; 373 C; 411 G; 482 T; 0 other;

Query Match 99.0%; Score 1746; DB 21; Length 1760;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 attatgtgaaactaccctgcgattctctgctgcaagagactcgaggctcgccctccacccc 60
Db 1 attatgtgaaactaccctgcgattctctgctgcaagagactcgaggctcgccctccacccc 60
Qy 61 agtcagccttccctggcgggtggtgaaagagactcgaggctcgcttccaaaagtgc 120
Db 61 agtcagccttccctggcgggtggtgaaagagactcgaggctcgcttccaaaagtgc 120
Qy 121 cgcctgagtgagcttaccctcagcagcaaatgagccttcttccttcctgctg 180
Db 121 cgcctgagtgagcttaccctcagcagcaaatgagccttcttccttcctgctg 180
Qy 181 acatctgcctggcggcagagacagggactcagcggaatccaaacctgagtagtaa 240
Db 181 acatctgcctggcggcagagacagggactcagcggaatccaaacctgagtagtaa 240
Qy 241 ttccagtttccagcaacaagaagacagagcagtagtacaagatccctcagcagatgagaatt 300
Db 241 ttccagtttccagcaacaagaagacagagcagtagtacaagatccctcagcagatgagaatt 300
Qy 301 attactgtctactaaggaagtagtattccacagcccaaggttctctactattccaaaga 360
Db 301 attactgtctactaaggaagtagtattccacagcccaaggttctctactattccaaaga 360
Qy 361 aatcagctcttggtatgagattagtagcagtagaggaataatgtatgatacaacttacg 420
Db 361 aatcagctcttggtatgagattagtagcagtagaggaataatgtatgatacaacttacg 420

Qy 421 ttgatgaaagatttgggttggaagaccagaaagatgacatatgacagtagtatttctgta 480
Db 421 ttgatgaaagatttgggttggaagaccagaaagatgacatatgacagtagtatttctgta 480
Qy 481 gaagttagaaacccagtgatgaaactatttaggcgctggtggttcttggtactgta 540
Db 481 gaagttagaaacccagtgatgaaactatttaggcgctggtggttcttggtactgta 540
Qy 541 ccaggaaaaacagatttctaaaggaaatcaaataggatgaaagatttctctgataat 600
Db 541 ccaggaaaaacagatttctaaaggaaatcaaataggatgaaagatttctctgataat 600
Qy 601 ttctcttctgaacacaggttctgcatccactacacattgtctatgcacacattccacaga 660
Db 601 ttctcttctgaacacaggttctgcatccactacacattgtctatgcacacattccacaga 660
Qy 661 gctgtgagcttcagtgctacccttcagcttctgcatccactgacctggactgtaataatgct 720
Db 661 gctgtgagcttcagtgctacccttcagcttctgcatccactgacctggactgtaataatgct 720
Qy 721 ataactccttttagtaccttggaaagaccttattcgatatcttgaaacagagatggcag 780
Db 721 ataactccttttagtaccttggaaagaccttattcgatatcttgaaacagagatggcag 780
Qy 781 ttgacttagaagatctatatagcccaactggcaacttcttggaagggtttgttttt 840
Db 781 ttgacttagaagatctatatagcccaactggcaacttcttggaagggtttgttttt 840
Qy 841 ggaagaaaatccagagtggtggaactgaaccttcaacagagagtaagattatcacgc 900
Db 841 ggaagaaaatccagagtggtggaactgaaccttcaacagagagtaagattatcacgc 900
Qy 901 tgcaacctcgttaacttctcagtgccatgaagggaagaaactaaagaaacccgataccatt 960
Db 901 tgcaacctcgttaacttctcagtgccatgaagggaagaaactaaagaaacccgataccatt 960
Qy 961 ttctggccaggttctcctcctgggttaaacgctgtgtgggaactgtgctgtctccac 1020
Db 961 ttctggccaggttctcctcctgggttaaacgctgtgtgggaactgtgctgtctccac 1020
Qy 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagttaataaaatccacagagctct 1080
Db 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagttaataaaatccacagagctct 1080
Qy 1081 cagtgagacccaagacgggtgtcaggggttgacacaaactcaaccacgctggccctg 1140
Db 1081 cagtgagacccaagacgggtgtcaggggttgacacaaactcaaccacgctggccctg 1140
Qy 1141 gacacacatgaggagtgactgtgtgagagggagacagagagatagccgcatcac 1200
Db 1141 gacacacatgaggagtgactgtgtgagagggagacagagagatagccgcatcac 1200
Qy 1201 caccagcagcttctccacagagctgtgcagtllyllycagtggtgcttattatagagaac 1260
Db 1201 caccagcagcttctccacagagctgtgcagtllyllycagtggtgcttattatagagaac 1260
Qy 1261 gtatggttatctccatcccttaactcctcagttgtgtgtcgaagaccttcaacttcagg 1320
Db 1261 gtatggttatctccatcccttaactcctcagttgtgtgtcgaagaccttcaacttcagg 1320
Qy 1321 atttacagtgattctgaaagagagagacatcaaacagaatttagagttgtgcaacagctc 1380
Db 1321 atttacagtgattctgaaagagagagacatcaaacagaatttagagttgtgcaacagctc 1380
Qy 1381 ttctgagagagccttaaaagacagagagaaaggtcttcaatcgtggaagaaaaattaaa 1440
Db 1381 ttctgagagagccttaaaagacagagagaaaggtcttcaatcgtggaagaaaaattaaa 1440
Qy 1441 tgtgtattataatagatcacacagctagttcagagttaccatgacgtattccactagct 1500
Db 1441 tgtgtattataatagatcacacagctagttcagagttaccatgacgtattccactagct 1500
Qy 1501 ggggtctgtatttcagttcttctcgatacgacggttagggtaagtacagggaaaaaac 1560
Db 1501 ggggtctgtatttcagttcttctcgatacgacggttagggtaagtacagggaaaaaac 1560

Db	1497	gggtcttgattcagttcttcgatacgcgcttagggtaatgtccagtaacaggaataaac	1555
QY	1561	tgtgcaagtgcacactgattccgttgcttgcttaactctaaagctccatgtccctgggc	1620
Db	1557	tgtgcaagtgcacactgattccgttgcttgcttaactctaaagctccatgtccctgggc	1616
QY	1621	ctaaatcgtatataaatcgtgatttttttttttttttttttttttttgcataattcacatatgtaa	1680
Db	1617	ctaaatcgtatataaatcgtgatttttttttttttttttttttttttgcataattcacatatgtaa	1676
QY	1681	accagaacattctatgtactacaaacctgggtttttaaaaggaactatgttgcctatgaat	1740
Db	1677	accagaacattctatgtactacaaacctgggtttttaaaaggaactatgttgcctatgaat	1736
QY	1741	taaactgtgctgctgatagga	1764
Db	1737	taaactgtgctgctgatagga	1760
RESULT 2			
AAA51498			
ID	AAA51498 standard; cDNA; 1760 BP.		
XX	AC AAA51498;		
XX	AC		
XX	XX		
DT	26-SEP-2000 (first entry)		
XX	XX		
DE	Human growth factor homologue, ZVEGF3, cDNA.		
XX	XX		
KW	Vascular endothelial growth factor; homologue; zvegfg3; CUB domain;		
KW	cysteine knot; platelet-derived growth factor; PDGF; neuropilin;		
KW	chromosome 4q28.3; cytostatic; anti-psoriatic; anti-inflammatory;		
KW	anti-diabetic; ophthalmological; anti-rheumatic; anti-arthritis;		
KW	vulnerable; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key		
FT	Location/Qualifiers		
FT	154..1191		
FT	/*tag= a		
FT	/product= ZVEGF3		
FT	154..195		
FT	/*tag= b		
XX	XX		
PN	WC2000034474-A2.		
XX	XX		
PD	15-JUN-2000.		
XX	XX		
PF	07-DEC-1999; 99WO-US28968.		
XX	XX		
PR	07-DEC-1998; 98US-0207120.		
PR	06-JUL-1999; 99US-0142576.		
PR	21-OCT-1999; 99US-0161653.		
PR	12-NOV-1999; 99US-0165255.		
XX	XX		
PA	(ZYMO) ZYMOGENETICS INC.		
XX	XX		
PI	Gao Z, Hart CE, Piddington CS, Sheppard PO, Shoemaker KE;		
PI	Gilbertson DG, West JW;		
XX	XX		
DR	WPI; 2000-423420/36.		
XX	P-PSDB; AAY96858.		
XX	XX		
PT	Novel zvegfg3 polypeptides and nucleotides encoding them useful for		
PT	stimulating growth of smooth muscle cells and fibroblasts comprising an		
PT	epitope bearing portion of a specific amino acid sequence		
XX	XX		
PS	Claim 29; Page 146-148; 173pp; English.		
XX	XX		
CC	This cDNA encodes a human vascular endothelial growth factor homologue		
CC	designated ZVEGF3. Polypeptides comprising an epitope-bearing portion		
CC	human or murine ZVEGF3 are claimed. The growth factors comprise a growth		

CC	factor domain and a CUB domain (generic sequence motifs are shown in
CC	AA96859 and AA96860). The growth factor domain is characterized by an
CC	arrangement of cysteine residues and beta-strands that is characteristic
CC	of the "cysteine knot" structure of the platelet-derived growth factor
CC	(PDGF) family. The CUB domain shows homology to CUB domains in
CC	neuropilins, human bone morphogenetic protein-1, porcine seminal plasma
CC	protein, bovine acidic seminal fluid protein and Xenopus laevis
CC	tollid-like protein. Structural analysis and homology predict that
CC	ZVEGF3 polypeptides complex with a second polypeptide to form multimeric
CC	proteins. The human zvegf3 gene has been mapped to chromosome 4q28.3.
CC	ZVEGF3 is useful for stimulating the growth of fibroblasts or smooth
CC	muscles cells, for activating cell surface PDGF-alpha receptor and for
CC	inhibiting PDGF-alpha receptor mediated cellular processes. ZVEGF3 is
CC	useful for regulating (post-development) organ growth, regeneration and
CC	maintenance, as well as tissue maintenance and repair processes. ZVEGF3
CC	antagonists are useful for treating cancer, rheumatoid arthritis,
CC	diabetic retinopathy, ischemic limb disease, peripheral vascular
CC	disease, myocardial ischemia, vascular intimal hyperplasia,
CC	atherosclerosis, wound healing, chronic liver disease and haemangioma
CC	formation. ZVEGF3 can also be used to modulate neurite growth and
CC	development of the nervous system, and for treating neurodegenerative
CC	diseases.
XX	
SQ	Sequence 1760 BP; 494 A; 373 C; 411 G; 482 T; 0 other;
	Query Match 99.0%; Score 1746; DB 21; Length 1760;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 1760; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY	1 attatgtggaactacacctcgattctctgtgcagagcaggctcggcgcttccacccc 60
DB	
QY	61 agtcagccttccctggcggttgaaagagactcgggagtgcgtcttccaaagtgc 120
DB	
QY	61 agtcagccttccctggcggttgaaagagactcgggagtgcgtcttccaaagtgc 120
DB	
QY	121 cgccgtgagtgcctccaccagtcagcaaatgagcctctcgggcttccctgctg 180
DB	
QY	121 cgccgtgagtgcctccaccagtcagcaaatgagcctctcgggcttccctgctg 180
DB	
QY	181 acatctgccctggccgcccagagacaggggactcaggcggaaatccaacctgagtagtaa 240
DB	
QY	181 acatctgccctggccgcccagagacaggggactcaggcggaaatccaacctgagtagtaa 240
DB	
QY	241 ttccagtttccagacaacaggaaacagaacggagtagtaaacatctcagcatgagagaatt 300
DB	
QY	241 ttccagtttccagacaacaggaaacagaacggagtagtaaacatctcagcatgagagaatt 300
DB	
QY	301 attactgtctactaatgaagtattcacagcccaagtttctcatacttatccaaga 360
DB	
QY	301 attactgtctactaatgaagtattcacagcccaagtttctcatacttatccaaga 360
DB	
QY	361 aatacggcttgttgatgagtagtagtcagtcagtagagaaaatgtatgatacaacttacg 420
DB	
QY	421 ttigtatgaagatttggccttgaaagccccgaagatgacatatgcaagtatgatatttcta 480
DB	
QY	421 ttigtatgaagatttggccttgaaagccccgaagatgacatatgcaagtatgatatttcta 480
DB	
QY	481 gaagttgagaacccaagtgtgaactatataggcgctgggtgtggtctctggtactgtta 540
DB	
QY	481 gaagttgagaacccaagtgtgaactatataggcgctgggtgtggtctctggtactgtta 540
DB	
QY	541 ccagggaaaaacagatttctaaaggaaatcaaattaggataagatttgtatctgtatgaatat 600
DB	
QY	541 ccagggaaaaacagatttctaaaggaaatcaaattaggataagatttgtatctgtatgaatat 600
DB	
QY	601 ttctctctgaacacagggttctgcattccatacaaatgtcatgccaaattcacagaa 660
DB	
QY	601 ttctctctgaacacagggttctgcattccatacaaatgtcatgccaaattcacagaa 660
DB	

QY	661	gctgtgagtccttcagtgctaaccccttcagcttttgccacttggacctgctttaataatgct	720
DB	661	gctgtgagtccttcagtgctaaccccttcagcttttgccacttggacctgctttaataatgct	720
QY	721	ataactgcctttagtagaccttgaagaccttattcgatatcttgaaacagagagatggcag	780
DB	721	ataactgcctttagtagaccttgaagaccttattcgatatcttgaaacagagagatggcag	780
QY	781	tggacttagaagatctatataggccaaacttgccaaacttcttgcaaggctttgttttt	840
DB	781	tggacttagaagatctatataggccaaacttgccaaacttcttgcaaggctttgttttt	840
QY	841	ggaagaaaatccaagatggtgatatcgaaacttctaaacagaggaggtaagattatacagc	900
DB	841	ggaagaaaatccaagatggtgatatcgaaacttctaaacagaggaggtaagattatacagc	900
QY	901	tgcacacctctgaacttctcagtggtccataagggaagaaactaaagagaacccgataccatt	960
DB	901	tgcacacctctgaacttctcagtggtccataagggaagaaactaaagagaacccgataccatt	960
QY	961	tcttgccacggttgctctcgtgttaaactgtggtggaaactgtgctgttgtctccac	1020
DB	961	tcttgccacggttgctctcgtgttaaactgtggtggaaactgtgctgttgtctccac	1020
QY	1021	aattgcaatgaattgctcaattgtgtcccaagcaaaagttaactaaaaataaccagaggtcctt	1080
DB	1021	aattgcaatgaattgctcaattgtgtcccaagcaaaagttaactaaaaataaccagaggtcctt	1080
QY	1081	cagttgagacaaaagaccggtgtcagggtattgacaaaactcaactcaacgacgtggccctg	1140
DB	1081	cagttgagacaaaagaccggtgtcagggtattgacaaaactcaactcaacgacgtggccctg	1140
QY	1141	gagcaccatgaggagtgtgactgtgtgtgcagaggagacacagaggatagccgcatcac	1200
DB	1141	gagcaccatgaggagtgtgactgtgtgtgcagaggagacacagaggatagccgcatcac	1200
QY	1201	caccagcagctcttgcccagaactgtgcaattlyllygcagttggctgattctattagagaac	1260
DB	1201	caccagcagctcttgcccagaactgtgcaattlyllygcagttggctgattctattagagaac	1260
QY	1261	gtatgcgttatctccactccttaactcagttgtgtgtcctcaagagacctttcatcttcagg	1320
DB	1257	gtatgcgttatctccactccttaactcagttgtgtgtcctcaagagacctttcatcttcagg	1316
QY	1321	atttcagtgcatcttgaaaggagagacatcaaacagaaattaggagtgtgtgcaacagctc	1380
DB	1317	atttcagtgcatcttgaaaggagagacatcaaacagaaattaggagtgtgtgcaacagctc	1376
QY	1381	ttttgagaggagcctaaaggacagagaaaaaggtcttcaatcgtgaaagaaaaattaaa	1440
DB	1377	ttttgagaggagcctaaaggacagagaaaaaggtcttcaatcgtgaaagaaaaattaaa	1436
QY	1441	tgtgtattaaatagataccacgactagtttcagagttaccatgtacagttatccactagct	1500
DB	1437	tgtgtattaaatagataccacgactagtttcagagttaccatgtacagttatccactagct	1496
QY	1501	gggtctgtatttcagttcttcgatacagcgttaggggttaaattgctcagttacaggaaaaaac	1560
DB	1497	gggtctgtatttcagttcttcgatacagcgttaggggttaaattgctcagttacaggaaaaaac	1556
QY	1561	tgtgcaagtggacacactgattccgtgtgcttgccttaactctaaagctccatgtcctgggc	1620
DB	1557	tgtgcaagtggacacactgattccgtgtgcttgccttaactctaaagctccatgtcctgggc	1616
QY	1621	ctaaaaatcgtataaaaatcggatttttttttttttttttttttttttgcataattcaatagttaa	1680
DB	1617	ctaaaaatcgtataaaaatcggatttttttttttttttttttttttttgcataattcaatagttaa	1676
QY	1681	accagaacattcttatgtactacaaacctggttttttaaaaggaaactatgttgtctatgaat	1740
DB	1677	accagaacattcttatgtactacaaacctggttttttaaaaggaaactatgttgtctatgaat	1736

QY	1741	taaactgtgtcgtcgtgatagga	1760
Db	1737	taaactgtgtcgtcgtgatagga	1760
RESULT 3			
AAAY1990			
ID	AAA71990	standard; cDNA; 2668 BP.	
AC			
XX	AAA71990;		
DF	19-JAN-2001	(first entry)	
XX			
DE	Human VEGF-X cDNA.		
XX			
KW	VEGF-X; vascular endothelial growth		
KW	antirheumatic; antiarthritic; anti-		
KW	angiogenesis regulator; vasculariz-		
KW	rheumatoid arthritis; diabetic reti-		
KW	tissue regeneration; tissue repair		
KW	venous sore; diabetic ulcer; burns		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifier	
FT	CDS	267..1304	
FT		/*Tag= a	
FT		/product= "VEGF-X"	
XX			
PN	WO200037641-A2.		
PD	29-JUN-2000.		
XX			
PF	21-DEC-1999; 99WO-US30503.		
XX			
PR	22-DEC-1998; 98GB-0028377.		
PR	18-MAR-1999; 99US-0124967.		
PR	08-NOV-1999; 99US-0164131.		
XX			
PA	(JANC) JANSSEN PHARM NV.		
XX			
PI	Gordon RD, Sprengel JJ, Yon JR,		
PI	Dhanaraj SN, Xu J;		
XX			
DR	WPI; 2000-442669/38.		
DR	P-PSDB; AAB10644.		
XX			
PT	New vascular endothelial growth fac-		
PT	preventing diseases associated with		
PT	such as cancer, rheumatoid arthritis		
XX			
PS	Disclosure; Fig 30B; 127pp; English		
XX			
CC	This invention describes a novel vasc-		
CC	(VEGF-X) protein (Ia) and its encod-		
CC	vulnary, cytostatic, antirheumat-		
CC	antidiabetic activity and acts as a		
CC	regulator. An antisense molecule o-		
CC	or preventing cancer, rheumatoid a-		
CC	retinopathy by inhibiting angiogen-		
CC	vascularization including formation		
CC	vessels, growth and development of		
CC	and tissue repair in a subject. The		
CC	for preparing medicaments for treat-		
CC	pressure sores, venous sores, diab-		
CC	skin graft growth, tissue repair, b-		
CC	tissue regeneration and organ repair		
CC	vascularization. This sequence encod-		
CC	in the method of the invention.		
XX			
SQ	Sequence 2668 BP; 780 A; 511 C; 567		

Sequence 2668 BP; 780 A; 511 C; 567 G; 810 T; 0 other;

Db	1134	aattgcaatgaatgtccaatgtgcccaagcaaatgtactcaaaaaaacaccacgaggctcctt	1193
Qy	1081	cagttgagaccacaaacacccggtgtcaggggattgcacaaatcactcaccaactgcccctg	1140
Db	1194	cagttgagaccacaaacacccggtgtcaggggattgcacaaatcactcaccaactgcccctg	1253
Qy	1141	gagcaccatgagagtgtaactgtgtgcagaggagcacagaggagatagccgcatacac	1200
Db	1254	gagcaccatgagagtgtaactgtgtgcagaggagcacagaggagatagccgcatacac	1313
Qy	1201	caaccagcagctctgccacagactgtgcagttlylrcagtgctgattctatttagagaac	1260
Db	1314	caccagcagctctgccacagactgtgcagttlylrcagtgctgattctatttagagaac	1369
Qy	1261	gtatcgcttattccatccttaactcagttgttgccttcaaggacgttctatcttcagg	1320
Db	1370	gtatcgcttattccatccttaactcagttgttgccttcaaggacgttctatcttcagg	1429
Qy	1321	atttacagtgcaattctgaagaggagacatcaaacagaattaggagttgtgcaacagctc	1380
Db	1430	atttacagtgcaattctgaagaggagacatcaaacagaattaggagttgtgcaacagctc	1489
Qy	1381	ttttgagaggagcctaaaggacagaggagaaaaaggtcttcaatcgctggaaagaaaaataa	1440
Db	1490	ttttgagaggagcctaaaggacagaggagaaaaaggtcttcaatcgctggaaagaaaaataa	1549
Qy	1441	tggtattataatagataccacagctagtttcagagttaccagttaactgtaacttaccactagct	1500
Db	1550	tggtattataatagataccacagctagtttcagagttaccagttaactgtaacttaccactagct	1609
Qy	1501	gggtctgtatttcaggtcttttcgatacggcttaggttaattgctcagtcacaggaaaaaac	1560
Db	1610	gggtctgtatttcaggtcttttcgatacggcttaggttaattgctcagtcacaggaaaaaac	1669
Qy	1561	tgtgcaagtgcacactgattccgttgcttcgttaacttaaaagctccatgtctctcgggc	1620
Db	1670	tgtgcaagtgcacactgattccgttgcttcgttaacttaaaagctccatgtctctcgggc	1729
Qy	1621	ctaaatcgtataaaatcggatt	1680
Db	1730	ctaaatcgtataaaatcggatt	1789
Qy	1681	accagaacattctatgtactacaaacctggttttttttttttttttttttttttttttttttt	1740
Db	1790	accagaacattctatgtactacaaacctggttttttttttttttttttttttttttttttttt	1849
Qy	1741	taaaactgtgctgctgatagga	1764
Db	1850	taaaactgtgctgctgatagga	1873
RESULT	4		
AAA47452			
ID	AAA47452	standard; cDNA; 2839 BP.	
XX			
AC	AAA47452;		
XX			
DT	20-OCT-2000	(first entry)	
XX			
DE	Human TANGO 128	coding sequence.	
XX			
KW	TANGO; 128; 140; 197; 212; 213; 224; 239;	modulating agent; asthma;	
KW	graft versus host diseases; rheumatoid arthritis; psoriasis;		
KW	inflammatory bowel disease; septic shock; ulcerative colitis;		
KW	Crohn's disease; chronic myelogenous leukemia; cancer; liver		
KW	disease; Hodgkin's disease; osteoarthritis; Lyme's disease;		
KW	cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;		
KW	systemic lupus erythematosus; transgenic animal; diagnosis;		
XX	prognosis; prophylactic; therapeutic; human; ds.		
OS	Homo sapiens.		
XX			

FH Key Location/Qualifiers
FT CDS 288..1325
FT /*tag= a
XX /product= TANGO 38
PN WO200039284-A1.
XX 06-JUL-2000.
PD
PF 23-DEC-1999; 99WO-US31025.
XX 30-DEC-1998; 98US-0223546.
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA Holtzman DA;
XX
XX WPI; 2000-465743/40.
DR P-PSDB; AAB01419.
XX
XX Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213,
PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases
XX
PS Claim 1; Fig 1; 209pp; English.
XX
XX Nucleic acids encoding TANGO polypeptides are useful as modulating
CC agents for regulating cellular processes like asthma, graft
CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
CC lupus erythematosus. The nucleic acids are also useful for producing
CC transgenic animals and the TANGO polypeptides themselves. Partial
CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
CC forensic biology, for diagnostic assays, prognostic assays,
CC pharmacogenomics and for monitoring clinical trials. TANGO
CC polypeptides are suitable for both prophylactic and therapeutic
CC methods for treating a subject at risk of a disorder or having a
CC disorder associated with aberrant TANGO expression. A wide range
CC of cellular disorders can be treated.
XX
SQ Sequence 2839 BP; 858 A; 532 C; 600 G; 849 T; 0 other;

Query Match 98.3%; Score 1734; DB 21; Length 2839;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1759; Conservative 0; Mismatches 0; Indels 5; Gaps 2;

Qy 1 attatgtgaaactaccctgcgattctctgtgccagagcaggctcgcgctccacccc 60
Dbb
Dbb 135 attatgtgaaactaccctgcgattctctgtgccagagcaggctcgcgctccacccc 194
Qy 61 agtcagccttccctgcggtggtgaaagagactcggagtcgctgtccaaagtgc 120
Dbb 195 agtcagccttccctgcggtggtgaaagagactcggagtcgctgtccaaagtgc 254
Qy 121 qcctgagtagctctaccctcagcgaatgagcctcttcgggtctctctctg 180
Dbb 255 qcctgagtagctctaccctcagcgaatgagcctcttcgggtctctctctg 314
Qy 181 acatctgcctgcgcccagagacagggactcagcggaatccaaacctgagtagtaa 240
Dbb 315 acatctgcctgcgcccagagacagggactcagcggaatccaaacctgagtagtaa 374
Qy 241 ttccagttttccagcaacaaggaacagacggagtagtacaagatctctcagcatgagaatt 300
Dbb 375 ttccagttttccagcaacaaggaacagacggagtagtacaagatctctcagcatgagaatt 434
Qy 301 attactgtctactaatggaagtattcacagcccaaggtttctcattatccaaaga 360
Dbb 435 attactgtctactaatggaagtattcacagcccaaggtttctcattatccaaaga 494

QY 361 aatacggctctgttatggagattagtagcagtagagagaaatgtatggtacaaactaacg 420
Dbb 495 aatacggctctgttatggagattagtagcagtagagagaaatgtatggtacaaactaacg 554
QY 421 ttgatgaaagatttggcttgaagaccagagatgacatatgcaagtattgatttcta 480
Dbb 555 ttgatgaaagatttggcttgaagaccagagatgacatatgcaagtattgatttcta 614
QY 481 gaagttaggaacccagtgatggaactatattagggcgctggtgtctctggtactgta 540
Dbb 615 gaagttaggaacccagtgatggaactatattagggcgctggtgtctctggtactgta 674
QY 541 ccagggaacacagatttctaaaggaaatcaaataggtaagatttgttatctgatgaatat 600
Dbb 675 ccagggaacacagatttctaaaggaaatcaaataggtaagatttgttatctgatgaatat 734
QY 601 ttctcttgaaacagggttctgcatccactacaacattgtcatgcccacaaatccacagaa 660
Dbb 735 ttctcttgaaacagggttctgcatccactacaacattgtcatgcccacaaatccacagaa 794
QY 661 gctgtaagctctcagtgctacccttcagcttgcactgacctgacctgcttaataatgct 720
Dbb 795 gctgtaagctctcagtgctacccttcagcttgcactgacctgacctgcttaataatgct 854
QY 721 ataactgcttttagtaaccttggaaagaccttattcgatatcttgaaccagagagatggcag 780
Dbb 855 ataactgcttttagtaaccttggaaagaccttattcgatatcttgaaccagagatggcag 914
QY 781 ttggactagaagatctatatagccaaactggcaacttcttggaagggtttgtttttt 840
Dbb 915 ttggactagaagatctatatagccaaactggcaacttcttggaagggtttgtttttt 974
QY 841 ggaagaaatccagagtggtggaatctgaacctctcaacagagaggttaagattatagc 900
Dbb 975 ggaagaaatccagagtggtggaatctgaacctctcaacagagaggttaagattatagc 1034
QY 901 tgcaacctgtaaccttctcagtgctcataagggaaacaaactaaagaaacccgataccatt 960
Dbb 1035 tgcaacctgtaaccttctcagtgctcataagggaaacaaactaaagaaacccgataccatt 1094
QY 961 ttctggccagggttctctctgtttaaaccgctgtgtgtgggaactgtgctgttctccac 1020
Dbb 1095 ttctggccagggttctctctgtttaaaccgctgtgtgtgggaactgtgctgttctccac 1154
QY 1021 aattgcaatgaatgcaatgtgtcccaagcaaatgtactaaaaataaccacgaggtcctt 1080
Dbb 1155 aattgcaatgaatgcaatgtgtcccaagcaaatgtactaaaaataaccacgaggtcctt 1214
QY 1081 cagttgagaccaaagaccggtgtcaggggattgcacaaatcactcaaccgctggccctg 1140
Dbb 1215 cagttgagaccaaagaccggtgtcaggggattgcacaaatcactcaaccgctggccctg 1274
QY 1141 gacacccatgagagagtgactgtgtgcagaggagacacagagagatagccgcatcac 1200
Dbb 1275 gacacccatgagagagtgactgtgtgcagaggagacacagagagatagccgcatcac 1334
QY 1201 caccagcagctctgccagagctgtgcagtllyllycagtggtgcttatttagagaac 1260
Dbb 1335 caccagcagctctgccagagctgtgtcagtl---gagtggtgcttatttagagaac 1390
QY 1261 gtagtggtatctccatcttaattctcagttgtttgttcaaggaccttctcttcagg 1320
Dbb 1391 gtagtggtatctccatcttaattctcagttgtttgttcaaggaccttctcttcagg 1450
QY 1321 attacagtgcatcttgaaagagagacatcaaacagaattagaggtgtgtgcaacagctc 1380
Dbb 1451 attacagtgcatcttgaaagagagacatcaaacagaattagaggtgtgtgcaacagctc 1510
QY 1381 ttgtgagagagccttaagagacagagaaaaggttctcattcgtggaagaaataaaa 1440
Dbb 1511 ttgtgagagagccttaagagacagagaaaaggttctcattcgtggaagaaataaaa 1570


```
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
XX
XX 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 31-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081209.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081953.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083332.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.

PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
XX (GETH ) GENENTECH INC.
XX
XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
PI
XX WPI; 1999-551358/46.
XX P-PSDB; AAY41766.
XX
XX New secreted and transmembrane polypeptides and their polynucleotides,
XX useful for treating blood coagulation disorders, cancers and cellular
XX adhesion disorders -
XX
XX Claim 2; Fig 206; 530pp; English.
XX
XX The present invention describes secreted and transmembrane polypeptides
XX and their polynucleotides. The nucleotide sequences are useful as
XX sources of probes, primers, for chromosome mapping, and for generation
XX of antisense sequences. They can also be used to create transgenic
XX animals. The proteins can be used to treat a variety of diseases and
XX disorders, depending on their function. Diseases that may be treated
XX include blood coagulation disorders, cancers and cellular adhesion
XX disorders. They may also be used to raise antibodies. AAZ33891 to
XX AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
XX polypeptide sequence given in the exemplification of the present
XX invention.
XX
XX Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;
SQ

Query Match 98.2%; Score 1732.4; DB 20; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtggaaactaccctgcgattctctgctgcagagcaggctcgcttccacccc 60
Db 132 attatgtggaaactaccctgcgattctctgctgcagagcaggctcgcttccacccc 191
QY 61 agtcagccttccctggcgggtggaagagactcggagctgcttccaaagtgc 120
Db 192 agtcagccttccctggcgggtggaagagactcggagctgcttccaaagtgc 251
QY 121 cgcctgagtgagctctcaccctcagccaaatgagctcttcgggtctctcctg 180
Db 252 cgcctgagtgagctctcaccctcagccaaatgagctcttcgggtctctcctg 311
QY 181 acatctgccttccctggcgggtggaagagactcggagctgcttccaaagtgc 240
Db 312 acatctgccttccctggcgggtggaagagactcggagctgcttccaaagtgc 371
QY 241 ttccagttttccagcaacaaggaacgagacgagtagtacaagatcctcagcagagatt 300
Db 372 ttccagttttccagcaacaaggaacgagacgagtagtacaagatcctcagcagagatt 431
```

[illegible]

PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX WPI; 2000-611443/58.
DR P-PSDB; AAB44322.
DR
XX
XX

PT Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities -
XX
XX

PS Claim 2; Fig 206; 636pp; English.
XX

XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting
CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX

SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.28; Score 1732.4; DB 21; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtggaacaccctcgattctctgctgcagagcaggtcgcgcttcacccc 60
DB 132 attatgtggaacaccctcgattctctgctgcagagcaggtcgcgcttcacccc 191
QY 61 agtgcagcttccctgctgctgaaagagactcggagctgcttcacaaagtcc 120
DB 192 agtgcagcttccctgctgctgaaagagactcggagctgcttcacaaagtcc 251
QY 121 cgcgtgagtgagctctcaccctcagtcagcaaatgagctcttcggtctctctg 180
DB 252 cgcgtgagtgagctctcaccctcagtcagcaaatgagctcttcggtctctctg 311
QY 181 acatctgctctgctcagcagagcagggactcagcggaatcaacctgagtagtaa 240
DB 312 acatctgctctgctcagcagagcagggactcagcggaatcaacctgagtagtaa 371
QY 241 ttccagttttccagcaacaaggaaacagagagtagtaacaagatcctcagtagagaatt 300
DB 372 ttccagttttccagcaacaaggaaacagagagtagtaacaagatcctcagtagagaatt 431
QY 301 attactgctctactaatgaaagtattcagagcccaagggttctcactatcacaaga 360
DB 432 attactgctctactaatgaaagtattcagagcccaagggttctcactatcacaaga 491
QY 361 aatcaggtctgtatggaatagtagcagtagcagaaataatgattgatacaacttacg 420
DB 492 aatcaggtctgtatggaatagtagcagtagcagaaataatgattgatacaacttacg 551
QY 421 ttgatgaaagatttgggttgaagcccaagagatgacatgacatgaaattgttga 480
DB 552 ttgatgaaagatttgggttgaagcccaagagatgacatgacatgaaattgttga 611
QY 481 gaagttaggaacccagtgatgaaactatatattagggcgctgggttctcgtactgta 540
DB 612 gaagttaggaacccagtgatgaaactatatattagggcgctgggttctcgtactgta 671
QY 541 ccaggaaaacagatttctaagaaatacaattaggaataagattgtatcgtgaatat 600
DB 672 ccaggaaaacagatttctaagaaatacaattaggaataagattgtatcgtgaatat 731
QY 601 ttctctctgaaaccaggttctgcatccactacacattgtcatgccacaattcacagaa 660

DB 732 ttctctctgaaaccaggttctgcatccactacaacattgtctgcacaaattcacagaa 791
QY 661 gctgtgagctcctcagtgctacccctcagcttttccactgagctgtcttaataatgct 720
DB 792 gctgtgagctcctcagtgctacccctcagcttttccactgagctgtcttaataatgct 851
QY 721 ataaactgccttagtaccttggaaagacctattcgatatcttgaaccagagatggcag 780
DB 852 ataaactgccttagtaccttggaaagacctattcgatatcttgaaccagagatggcag 911
QY 781 ttggacttagaagatctatataggccaacttggcaacttcttggcaaggctttgtttt 840
DB 912 ttggacttagaagatctatataggccaacttggcaacttcttggcaaggctttgtttt 971
QY 841 ggaagaaaaatccagagtggtgatctgaaccttctaacagagaggtgaagattatacagc 900
DB 972 ggaagaaaaatccagagtggtgatctgaaccttctaacagagaggtgaagattatacagc 1031
QY 901 tgcaacaccttgtaacttctcagtgctcataagggagaactaaagaaacccagataccatt 960
DB 1032 tgcaacaccttgtaacttctcagtgctcataagggagaactaaagaaacccagataccatt 1091
QY 961 ttctggccaggttgcctcctggttaaacgctgtggtggaaactgtgcctgtgtctccac 1020
DB 1092 ttctggccaggttgcctcctggttaaacgctgtggtggaaactgtgcctgtgtctccac 1151
QY 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagtactaaaaataccacaggtccctt 1080
DB 1152 aattgcaatgaatgcaatgtgtcccaagcaaaagtactaaaaataccacaggtccctt 1211
QY 1081 cagttgagaccaaagaccggtgtcaggggatgcaaaatcactccacgagctggccctg 1140
DB 1212 cagttgagaccaaagaccggtgtcaggggatgcaaaatcactccacgagctggccctg 1271
QY 1141 gagcaccatgagagtgactgtgtgctgagagggagcacagggagtagccgcatacc 1200
DB 1272 gagcaccatgagagtgactgtgtgctgagagggagcacagggagtagccgcatacc 1331
QY 1201 caaccagcagcttggcccaagagctgtcaggtlyllycagtggtgattctattagagaac 1260
DB 1332 caccagcagcttggcccaagagctgtcaggt---gcagtggtgattctattagagaac 1387
QY 1261 gtagcgttatctccatccttaactcagttgttcttcaagagaccttctcattccagg 1320
DB 1388 gtagcgttatctccatccttaactcagttgttcttcaagagaccttctcattccagg 1447
QY 1321 attacagtgcatcttgaagaggagacatcaaacagaattaggagttgtgcaacagctc 1380
DB 1448 attacagtgcatcttgaagaggagacatcaaacagaattaggagttgtgcaacagctc 1507
QY 1381 ttttgagaggagagcctaaaggacagagaaaagggtcttcaatctggaagaaaataaa 1440
DB 1508 ttttgagaggagagcctaaaggacagagaaaagggtcttcaatctggaagaaaataaa 1567
QY 1441 tttgtattaaatagatcacaccagctagttcagagttaccatgtacgtattccactagct 1500
DB 1568 tttgtattaaatagatcacaccagctagttcagagttaccatgtacgtattccactagct 1627
QY 1501 ggggtctgtattctcagttcttctgatacggcttaggtaattgagtagcagagaaaaaac 1560
DB 1628 ggggtctgtattctcagttcttctgatacggcttaggtaattgagtagcagagaaaaaac 1687
QY 1561 tftgcaagtgaacacactgattccgttgccttgaactcctaaagctccatgtcctcgggc 1620
DB 1698 tftgcaagtgaacacactgattccgttgccttgaactcctaaagctccatgtcctcgggc 1747
QY 1621 ctaaaaatcgataaaaaatctggatttttttttttttttttttttttttttttttttttt 1680
DB 1748 ctaaaaatcgataaaaaatctggga-ttttttttttttttttttttttttttttttttttttt 1806
QY 1681 accagaacattctatgtactacaacacctgggtttttaaaggaactatgttgcctatgaat 1740

Db 1807 accagaaattctatgtactacaaacctggtttttaaaaggaactatgttgctatgaat 1866

QY 1741 taaactgtgtcgtgctgatagga 1764
|||||

Db 1867 taaactgtgtcgtgctgatagga 1890
|||||

RESULT 9
AAA88515
ID AAA88515 standard; cDNA; 2849 BP.
XX
AC AAA88515;
XX
XX 22-JAN-2001 (first entry)
XX
XX Human PRO200 (VEGF-E) cDNA clone DNA29101-1272.
XX
KW PRO200; vascular epithelial growth factor E; VEGF-E; human;
KW ocular disease; retinopathy; maculopathy; therapy;
KW retinitis pigmentosa; macular degeneration; retinal detachment;
KW retinal tear; macular hole; myopia; traumatic choriorretinopathy;
KW acute retinal necrosis syndrome; contusion; edema;
KW retinal vision occlusion; vascular disease; retinal vasculitis;
KW thrombocytopenic purpura; uveitis; retinal occlusion; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 285..1322
FT /*tag= a
FT sig_peptide 285..326
FT /*tag= b
FT mat_peptide 327..1319
FT /*tag= c
XX
XX W0200053760-A2.
XX
XX 14-SEP-2000.
XX
XX 10-MAR-2000; 2000WO-US06319.
XX
XX 12-MAR-1999; 9905-0123957.
XX
XX (GETH) GENENTECH INC.
XX
XX Ferrara N, Goddard A, Gurney AL, Hebert C, Henzel WJ, Kabakoff RC;
PI Klein RD, Kijavini IJ, Kuo SS, La Fleur M, Wood WI;
XX
XX WPI; 2000-587437/55.
DR P-PSDB; AAB19578.
XX
XX Novel PRO polypeptides useful for preventing or rescuing retinal cells
PT from injury caused by ocular diseases such as retinitis pigmentosa,
PT retinopathy, retinal degenerative diseases, degenerative myopia,
PT uveitis -
XX
XX Example 8; Fig 1; 140pp; English.
PS
XX
XX The present sequence is that of cDNA clone DNA29101-1272 or UNQ174
CC (ATCC 209653) isolated from a human glioma cell line G61 cDNA
CC library using probes (see AAA88523-26) based on an expressed sequence
CC tag (see AAA88522) identified on the basis of homology to vascular
CC endothelial growth factor (VEGF). The clone contains a single open
CC reading frame encoding a 345-amino acid protein (see AAB19578)
CC designated PRO200 or VEGF-E. The isolated cDNA can be used in the
CC recombinant production of PRO200 (VEGF-E). The invention relates
CC to the use of PRO polypeptides, including PRO200, to delay, prevent
CC or rescue retinal cells such as retinal neurons selected from
CC photoreceptors, retinal ganglion cells, displaced retinal ganglion
CC cells, amacrine cells, displaced amacrine cells, horizontal and
CC bipolar neurons, and supportive cells (including Mueller cells and
CC pigment epithelial cells) from injury and degradation. The retinal
CC cells are preferably photoreceptors and photoreceptor cell injury or

CC death is caused by retinal injury, light or environmental trauma or
CC by an ocular disease selected from retinitis pigmentosa, macular
CC degeneration, including age-related, retinal detachment, retinal
CC tears, retinopathy, retinal degenerative diseases, macular holes,
CC degenerative myopia, acute retinal necrosis syndrome, traumatic
CC choriorretinopathies or contusion such as Purtscher's retinopathy,
CC edema, ischemic conditions such as central or branch retinal vision
CC occlusion, collagen vascular diseases, thrombocytopenic purpura,
CC uveitis, retinal vasculitis and occlusion associated with Eales
CC disease and systemic lupus erythematosus (claimed).
XX
XX Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;
SQ

Query Match 98.2%; Score 1732.4; DB 21; Length 2849;
Best Local Similarity 99.7%; Pred No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtgaaactaccctgcgattctctgtcgcagagcagctcgcgttcacccc 60
|||||
Db 132 attatgtgaaactaccctgcgattctctgtcgcagagcagctcgcgttcacccc 191
|||||

QY 61 agtcgagccttcacctggcggtggtgaaagacactcggagctcgtctccaaagtgc 120
|||||
Db 192 agtcgagccttcacctggcggtggtgaaagacactcggagctcgtctccaaagtgc 251
|||||

QY 121 cgcctgagtgagctctcaccctcagcacaatgagctctcggcttcctcgtcg 180
|||||
Db 252 cgcctgagtgagctctcaccctcagcacaatgagctctcggcttcctcgtcg 311
|||||

QY 181 acatctgcctcgcgcgcagagacaggggactcagcgcgaatccaaactgagtgtaa 240
|||||
Db 312 acatctgcctcgcgcgcagagacaggggactcagcgcgaatccaaactgagtgtaa 371
|||||

QY 241 ttccagtttccagcaacaaggaacagacgaggtacacagatcctcagcatgagagaatt 300
|||||
Db 372 ttccagtttccagcaacaaggaacagacgaggtacacagatcctcagcatgagagaatt 431
|||||

QY 301 attactgtgtactaataatggaagtattcagaccccaagggttccctcactatccaaga 360
|||||
Db 432 attactgtgtactaataatggaagtattcagaccccaagggttccctcactatccaaga 491
|||||

QY 361 aatacgcgttcgtgtatggagattagtagcagtagaggaataatgtaggatacaacttacy 420
|||||
Db 492 aatacgcgttcgtgtatggagattagtagcagtagaggaataatgtaggatacaacttacy 551
|||||

QY 421 ttgtatgaagaatttggcgttgaaagaccagacagatgacatgcaagtattgatttga 480
|||||
Db 552 ttgtatgaagaatttggcgttgaaagaccagacagatgacatgcaagtattgatttga 611
|||||

QY 481 gaagttgaggaaccagtgatgaaactattagggcgtggtgttctggtactgtga 540
|||||
Db 612 gaagttgaggaaccagtgatgaaactattagggcgtggtgttctggtactgtga 671
|||||

QY 541 ccaggaaaaacagattcttaaaagaaaatacgaataagagattgttctctgataaat 600
|||||
Db 672 ccaggaaaaacagattcttaaaagaaaatacgaataagagattgttctctgataaat 731
|||||

QY 601 ttctctctgaaccagggttctgtcactcacaacattgtcatgccacaattccacagaa 660
|||||
Db 732 ttctctctgaaccagggttctgtcactcacaacattgtcatgccacaattccacagaa 791
|||||

QY 661 gctgtgagtcctcagtgctaccccttcagcttgcactggagcctgcttaataatgct 720
|||||
Db 792 gctgtgagtcctcagtgctaccccttcagcttgcactggagcctgcttaataatgct 851
|||||

QY 721 ataatgccttttagaccttggaagaccttattcgatatcttgaaccagagatggcag 780
|||||
Db 852 ataatgccttttagaccttggaagaccttattcgatatcttgaaccagagatggcag 911
|||||

QY 781 ttggactagaagatctatagggcacaacttggcacaacttcttggcaagcctttgtttt 840
|||||
Db 912 ttggactagaagatctatagggcacaacttggcacaacttcttggcaagcctttgtttt 971
|||||

P-PSDB: AAB33414.

Sixty four PRO polypeptides, useful in the diagnosis and treatment of immune related disorders, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, thyroiditis and diabetes mellitus -

Claim 23; Fig 1; 309pp; English.

The present invention describes sixty four human PRO proteins which can be used in the treatment of immune related diseases. The human PRO proteins, anti-PRO antibodies, agonists and antagonists are useful for treating and diagnosing immune related disorders. The disorders are selected from systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, spondyloarthropathies, systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central and peripheral nervous systems, hepatobiliary diseases, inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's disease, autoimmune or immune-mediated skin diseases, allergic diseases, immunological diseases of the lung, and transplantation associated diseases including graft rejection and graft-versus-host-disease. AAC58397 to AAC58578 represent PCR primers and hybridisation probes used in the isolation of human PRO sequences. AAC58579 to AAC58642 and AAB33414 to AAB33477 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention.

Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 21; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY	1	attatgtggaatacactcgtcgtattctctgtccagagcaggctcgcgtctccacccc	60
Db	132	attatgtggaatacactcgtcgtattctctgtccagagcaggctcgcgtctccacccc	191
QY	61	agtgcagcttcccttgcggtgtaaaagagactcgggagtcgtcttccaaatgccc	120
Db	192	agtgcagcttcccttgcggtgtaaaagagactcgggagtcgtcttccaaatgccc	251
QY	121	cgccgtgagtgagctctcaccacagtcagccaaatgagcctctcgggtcttccctgctg	180
Db	252	cgccgtgagtgagctctcaccacagtcagccaaatgagcctctcgggtcttccctgctg	311
QY	181	acatctgccctggccggccagagacacgtgactcaggcggaatcccaactgagtagtaa	240
Db	312	acatctgccctggccggccagagacacgtgactcaggcggaatcccaactgagtagtaa	371
QY	241	ttccagtttccagcaacaaggacagacaggagtagtacaagatcctcagcatgagagaatt	300
Db	372	ttccagtttccagcaacaaggacagacaggagtagtacaagatcctcagcatgagagaatt	431
QY	301	attactgtgtactaatggaagatattcacagcccaaggtttctctcatactatccaaga	360
Db	432	attactgtgtactaatggaagatattcacagcccaaggtttctctcatactatccaaga	491
QY	361	aatacggctcttgtagtgagattagtagcagtagagagaaaatgtatggatacaacttaag	420
Db	492	aatacggctcttgtagtgagattagtagcagtagagagaaaatgtatggatacaacttaag	551
QY	421	tttgatgaagatttggcttgaagaccagacagatgcacatatgaagtagtatgtttgtga	480
Db	552	tttgatgaagatttggcttgaagaccagacagatgcacatatgaagtagtatgtttgtga	611
QY	481	gaagttgaggaaccagctgatgaaactatattagggcgtgggtggttctggtacttgta	540
Db	612	gaagttgaggaaccagctgatgaaactatattagggcgtgggtggttctggtacttgta	671
QY	541	ccaggaaaacagatttctaaagaaatcaaataggataagatgtttgtatctgatgaatat	600

QY 1681 accagaaacatttatgtactacaaacctggtttttaaaaggaaactatgttgcataat 1740
|||||
Db 1807 accagaaacatttatgtactacaaacctggtttttaaaaggaaactatgttgcataat 1866
|||||
QY 1741 taaactgtgctgctgataaga 1764
|||||
Db 1867 taaactgtgctgctgataaga 1890
|||||

RESULT 11

AAA77621
ID AAA77621 standard; cDNA; 2849 BP.

XX
AC AAA77621;

XX 07-NOV-2000 (first entry)

XX Human PRO713 cDNA sequence SEQ ID NO:136.

XX Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;
KW cytosolic; gene therapy; vaccine; ss.

XX Homo sapiens.

XX WO200032221-A2.

XX 08-JUN-2000.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1998; 98WO-US25108.

PR 16-DEC-1998; 98US-0112850.

PR 12-JAN-1999; 99US-0115554.

PR 08-MAR-1999; 99WO-US05028.

PR 12-MAR-1999; 99US-0123957.

PR 28-APR-1999; 99US-0131445.

PR 14-MAY-1999; 99US-0134287.

PR 02-JUN-1999; 99WO-US12252.

PR 23-JUN-1999; 99US-0141037.

PR 20-JUL-1999; 99US-0144758.

PR 26-JUL-1999; 99US-0145698.

PR 01-SEP-1999; 99WO-US20111.

PR 08-SEP-1999; 99WO-US20594.

PR 13-SEP-1999; 99WO-US20944.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 05-OCT-1999; 99WO-US23089.

PR 29-OCT-1999; 99US-0162506.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;

PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;

PI Watanabe CK, Williams PM, Wood WI;

XX WPI; 2000-412154/35.

CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate PRO expression such as cardiovascular, endothelial or
CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAA77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.

XX
SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 21; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtgaaactacccctgcgattctctctgctccagagcaggctcgccctccacc 60
|||||
Db 132 attatgtgaaactacccctgcgattctctctgctccagagcaggctcgccctccacc 191
|||||
QY 61 agtcagcctccctcgcggtggtgaaagactcgggagctgctctccaaagtgc 120
|||||
Db 192 agtcagcctccctcgcggtggtgaaagactcgggagctgctctccaaagtgc 251
|||||
QY 121 cgcgtgagtgagctctcaccctcagtcagccaaatgagcctctcgggctctcctgctg 180
|||||
Db 252 cgcgtgagtgagctctcaccctcagtcagccaaatgagcctctcgggctctcctgctg 311
|||||
QY 181 acatctgctcctcgccgagagacagggactcagcggaatcccaacctgagtagtaa 240
|||||
Db 312 acatctgctcctcgccgagagacagggactcagcggaatcccaacctgagtagtaa 371
|||||
QY 241 ttccagtttccagcaacaaaggaacgagacgagatcaagatctcctcagcagagaa 300
|||||
Db 372 ttccagtttccagcaacaaaggaacgagacgagatcaagatctcctcagcagagaa 431
|||||
QY 301 attactgtctactaaatggaagtattcacagcccaagggttctcctacatacttccaa 360
|||||
Db 432 attactgtctactaaatggaagtattcacagcccaagggttctcctacatacttccaa 491
|||||
QY 361 aatcggctctgtagtgagattagtagcagtagagagaaaatgtagtacaaacttacg 420
|||||
Db 492 aatcggctctgtagtgagattagtagcagtagagagaaaatgtagtacaaacttacg 551
|||||
QY 421 ttgatgaaagatttggtctgaaagacccagagatgacatagcaagtattgttga 480
|||||
Db 552 ttgatgaaagatttggtctgaaagacccagagatgacatagcaagtattgttga 611
|||||
QY 481 gaagttgaggaacccagtgatgaaactatattaggcgctggtgtgtctgtactgtga 540
|||||
Db 612 gaagttgaggaacccagtgatgaaactatattaggcgctggtgtgtctgtactgtga 671
|||||
QY 541 ccaggaaaacagatttttaaggaaatcaaataggataagattgttatgtatgaatat 600
|||||
Db 672 ccaggaaaacagatttttaaggaaatcaaataggataagattgttatgtatgaatat 731
|||||
QY 601 ttctcttgaacacaggttctgcctccactacaacattgtcatgcacaaatccacagaa 660
|||||
Db 732 ttctcttgaacacaggttctgcctccactacaacattgtcatgcacaaatccacagaa 791
|||||
QY 661 gctgtgagctcttcagtgctaccctctcagctttgcccactggacctgcttaataatgct 720
|||||
Db 792 gctgtgagctcttcagtgctaccctctcagctttgcccactggacctgcttaataatgct 851
|||||
QY 721 ataactgccttttagtaccttggaagaccttattgatatttgaaccagagatggcag 780
|||||
Db 852 ataactgccttttagtaccttggaagaccttattgatatttgaaccagagatggcag 911
|||||
QY 781 ttgaccttaagaactctatatagccaacttgcaactctcttgcaaggctttgttttt 840
|||||
Db 912 ttgaccttaagaactctatatagccaacttgcaactctcttgcaaggctttgttttt 971
|||||
QY 841 ggaagaaaatccagagtggtggatctgaaccttctaacagaggggtaagattatacagc 900
|||||

Db 972 ggaagaaacccagagtggtggaactctgaacctctaaacagagagtggaagtattacagc 1031
QY 901 tgcacacctgtaactctcagttcagttcacaaggaagaaactaaagagacacgataccatt 960
Db 1032 tgcacacctgtaactctcagttcagttcacaaggaagaaactaaagagacacgataccatt 1091
QY 961 ttctggccaggttgctctcctgggttaaacgctgtggtgggaactgtgctgttgcctccac 1020
Db 1092 ttctggccaggttgctctcctgggttaaacgctgtggtgggaactgtgctgttgcctccac 1151
QY 1021 aattgcaatgaatgctcaatgtgtcccaagcaaatgtaataaaataccacagagtcctt 1080
Db 1152 aattgcaatgaatgctcaatgtgtcccaagcaaatgtaataaaataccacagagtcctt 1211
QY 1081 cagttagacacaaagacgcgtgtcaggggattgcacaaatcactcacccagcgtgcccctg 1140
Db 1212 cagttagacacaaagacgcgtgtcaggggattgcacaaatcactcacccagcgtgcccctg 1271
QY 1141 gacacacatgagagtgactgtgtgcagagggagacagagagtagccgcacac 1200
Db 1272 gacacacatgagagtgactgtgtgcagagggagacagagagtagccgcacac 1331
QY 1201 caccagcagctctgccacagcgtgtgcagtllylgycaagtgctgattctattagagac 1260
Db 1332 caccagcagctctgccacagcgtgtgcagtllylgycaagtgctgattctattagagac 1387
QY 1261 gtatggttattccatcccttaactcagttgtgttgccttcagagacatttcatctcagg 1320
Db 1388 gtatggttattccatcccttaactcagttgtgttgccttcagagacatttcatctcagg 1447
QY 1321 attacagtgattctgaagagagagacatcaacagagattagagttgtgcaacagctc 1380
Db 1448 attacagtgattctgaagagagagacatcaacagagattagagttgtgcaacagctc 1507
QY 1381 ttttgagagagggcctaagagacagagagaaaggtcttcaatcgtggaaagaaattaaa 1440
Db 1508 ttttgagagagggcctaagagacagagagaaaggtcttcaatcgtggaaagaaattaaa 1567
QY 1441 tgtgtattataatagatcacacagtagtttccagagttaccatgtacgtattccactagct 1500
Db 1568 tgtgtattataatagatcacacagtagtttccagagttaccatgtacgtattccactagct 1627
QY 1501 ggtctgtatttcagttcttcagacgcttagggttaagttaagtcagtcagacagaaaaaac 1560
Db 1628 ggtctgtatttcagttcttcagacgcttagggttaagttaagtcagtcagacagaaaaaac 1687
QY 1561 tgtgcaagtgcacacctgattccgttgcctgttaactctaaagctccatgtcctgggc 1620
Db 1688 tgtgcaagtgcacacctgattccgttgcctgttaactctaaagctccatgtcctgggc 1747
QY 1621 ctataatcgtataaaatctggatt 1680
Db 1748 ctataatcgtataaaatctgga-ttt 1806
QY 1681 accagacattctatgtactacaacacctggttttaaaaggaactatgttgcctatgaat 1740
Db 1807 accagacattctatgtactacaacacctggttttaaaaggaactatgttgcctatgaat 1866
QY 1741 taaactgtgtcgtgctgatagga 1764
Db 1867 taaactgtgtcgtgctgatagga 1890

RESULT 12
AAC90564
ID AAC90564 standard; cDNA; 2849 BP.
XX
AC AAC90564;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human PRO200 cDNA.

XX Human; PRO; cardiac; antiangiogenic; antiarteriosclerotic; hypotensive;
KW vasotropic; antirheumatic; antiarthritic; antiinflammatory; cytostatic;
KW vulnery; antiangular; gene therapy; cardiovascular disease;
KW endothelial disorder; angiogenic disorder; cancer; periodontal disease;
wound healing; ss.
XX Homo sapiens.
OS
XX WO200073445-A2.
XX
PD 07-DEC-2000.
XX
PF 17-MAY-2000; 2000WO-US13705.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0144758.
PR 28-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99US-0146222.
PR 30-NOV-1999; 99WO-US20111.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Gurney AL, Kuo SS, Mark MR, Marsters SA;
PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;
XX
DR WPI: 2001-025251/03.
DR P-PSDB; AAB50952.
XX
XX Seventeen nucleic acids encoding PRO polypeptides which are useful in
PT diagnosis and treatment of cardiovascular, endothelial or angiogenic
PT disorders in a mammal -
XX
PS Claim 60; Fig 3; 182pp; English.
XX
CC The present sequence is one of seventeen nucleic acids encoding PRO
CC polypeptides. The PRO nucleic acids, polypeptides, agonists and
CC antagonists are useful for treating cardiovascular, endothelial or
CC angiogenic disorders in a mammal. Examples of these disorders include
CC cardiac hypertrophy, trauma, cancer, age-related macular degeneration,
CC atherosclerosis, hypertension, arterial restenosis, Reynaud's disease,
CC rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis
CC and lymphangitis. The PRO polypeptides and antagonists are also used to
CC prevent tumour angiogenesis and for treating periodontal diseases. They
CC are also used to stimulate wound healing and tissue regeneration.
CC The PRO nucleic acids, polypeptides and anti-PRO antibodies are useful
CC for diagnosing a cardiovascular, endothelial or angiogenic disorder.
XX
SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 22; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
QY 1 attatgtggaactaccctgcgtattctctgctgcagagcagcgtcgcgtccacccc 60
|||||
DB 132 attatgtggaactaccctgcgtattctctgctgcagagcagcgtcgcgtccacccc 191

XX WPI: 2001-025022/03.
DR P-PSDB; AAB49895.
XX
PT
PT
PT
PT
PT
PT
XX
XX
PS
PS
PS
XX
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
XX
SQ

New compositions containing a PRO526, PRO719, PRO725, PRO1031 or PRO200 proteins for modulating immune response or proliferation of T-lymphocytes in mammal, especially for treating immune related disorders, e.g. graft rejection -

Claim 21; Fig 9; 133pp: English.

The present invention discloses the coding and protein sequences of human proteins PRO526, PRO719, PRO725, PRO1031 and PRO200. These proteins, of their coding sequences and antibodies can be used in the treatment of immune-related diseases, including systemic lupus erythematosus, rheumatoid arthritis, thyroiditis, immune-mediated renal disease, demyelinating diseases such as multiple sclerosis, hepatobiliary diseases including primary biliary cirrhosis, inflammatory bowel disease, immune-mediated skin diseases such as psoriasis, allergic diseases including asthma, immunologic diseases of the lung, transplantation associated diseases and infectious diseases such as HIV and hepatitis.

Sequence 2849 BP: 851 A; 528 C; 619 G; 850 T; 1 other:

Query Match	98.28;	Score 1732.4;	DB 22;	Length 2849;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1758;	Conservative 0;	Mismatches 1;	Indels 5;	Gaps 2;
QY 1	attatgtggaaactaccctcgattctctgtcgccagagcaggctcgcgcttcaccccc	60		
Db 132	attatgtggaaactaccctcgattctctgtcgccagagcaggctcgcgcttcaccccc	191		
QY 61	agtcagccttccccctggcggtgtgtgaagagactcgggagtcgtcttccaaagtgc	120		
Db 192	agtcagccttccccctggcggtgtgtgaagagactcgggagtcgtcttccaaagtgc	251		
QY 121	cgccgtgagtgagcttcacccagtcagccaaatgagccttcgggcttctcctgtg	180		
Db 252	cgccgtgagtgagcttcacccagtcagccaaatgagccttcgggcttctcctgtg	311		
QY 181	acatctgccttggcgcagagacaggggactcaggcggaaatccaaacctgagtagtaa	240		
Db 312	acatctgccttggcgcagagacaggggactcaggcggaaatccaaacctgagtagtaa	371		
QY 241	ttcagttttccagcaacaagaaacagacagagtagtacaagatcctcagatgagagaatt	300		
Db 372	ttcagttttccagcaacaagaaacagacagagtagtacaagatcctcagatgagagaatt	431		
QY 301	attactgtctactaatggaagtattcacagcccaaggtttcctcactattccaaaga	360		
Db 432	attactgtctactaatggaagtattcacagcccaaggtttcctcactattccaaaga	491		
QY 361	aatacgcttcttggtatggagattagtagcagtagaggaataatgtatgatacaacttacg	420		
Db 492	aatacgcttcttggtatggagattagtagcagtagaggaataatgtatgatacaacttacg	551		
QY 421	tttgatgaagatttgggcttgaagaccagcaagatgacatgatacaagtatgattttga	480		
Db 552	tttgatgaagatttgggcttgaagaccagcaagatgacatgatacaagtatgattttga	611		
QY 481	gaagttgagaacccagtgatggaactataataggcgctgggtctcctcgggtactgta	540		
Db 612	gaagttgagaacccagtgatggaactataataggcgctgggtctcctcgggtactgta	671		
QY 541	ccagggaacacagatttctaaaggaaatacaaataggataaagatttgcattctgcataat	600		
Db 672	ccagggaacacagatttctaaaggaaatacaaataggataaagatttgcattctgcataat	731		
QY 601	tttctctgaaacacaggttctgcattccactacaacattgtcatgccacaattcaacaga	660		
Db 732	tttctctgaaacacaggttctgcattccactacaacattgtcatgccacaattcaacaga	791		

Db 1867 taaactgtctatctgtatagga 1890
 RESULT 14
 AAC97404
 ID AAC97404 standard; cDNA; 2849 BP.
 AC AAC97404;
 XX
 DT 28-FEB-2001 (first entry)
 XX
 DE Human angiogenesis-associated protein PRO200 cDNA, SEQ ID NO:50.
 XX
 KW Human; angiogenesis-associated protein; PRO; endothelial cell growth;
 KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
 KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
 KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
 KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
 KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
 KW gene therapy; transgenic animal; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200053753-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 05-JAN-2000; 2000WO-US00219.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 12-MAR-1999; 99US-0123957.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144756.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;
 PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;
 PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WT;
 XX
 DR WPI; 2001-090793/10.
 DR P-PSDB; AAB53074.
 XX
 PT New isolated nucleic acid for producing a PRO polypeptide, analyzing
 PT genetic disorders and treating cardiovascular, endothelial or
 PT angiogenic disorders, such as atherosclerosis, wounds or cancer -
 XX
 XX Claim 58; Fig 21A-B; 293pp; English.
 XX
 CC The invention relates to novel human angiogenesis-associated proteins
 CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
 CC PRO proteins. The invention also relates to vectors and host cells
 CC comprising a PRO nucleic acid, the recombinant production of a PRO
 CC protein, PRO antibodies specific for a PRO protein, fusion proteins
 CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
 CC compounds which inhibit the expression of a PRO gene. The invention
 CC additionally encompasses methods of identifying modulators of PRO
 CC expression or activity; diagnosing a cardiovascular, endothelial or
 CC angiogenic disorder, or a susceptibility to such a disorder by detecting
 CC mutations in a PRO gene, or the expression level of a PRO gene within a

CC particular tissue; treating a cardiovascular, endothelial or angiogenic
 CC disorder via the administration of a PRO protein, PRO nucleic acid, or
 CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a
 CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial
 CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
 CC administration of a PRO protein, or an agonist or antagonist thereof.
 CC PRO nucleic acids, PRO proteins, or antibodies against PRO proteins, PRO
 CC agonists and PRO antagonists may be used as therapeutic agents to treat
 CC cardiovascular, endothelial or angiogenic disorders, such as
 CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,
 CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,
 CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
 CC disease, or stroke. PRO nucleic acids are additionally useful in the
 CC recombinant production of PRO proteins, as hybridisation probes to
 CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,
 CC to map genes encoding PRO proteins, to analyse genetic disorders, and in
 CC gene therapy. PRO nucleic acids can also be used to produce transgenic
 CC animals useful for the development and screening of potential
 CC therapeutic agents. The present sequence represents a cDNA encoding a PRO
 CC protein of the invention.
 XX
 SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 22; Length 2849;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
 QY 1 attatgtgaaactaccctgctgattctctgtcgcagagcagcgtcgcccttcacccc 60
 Db 132 attatgtgaaactaccctgctgattctctgtcgcagagcagcgtcgcccttcacccc 191
 QY 61 agtcagcctccctggcggtgtgaaagagactcggagtcgctctccaaagtgc 120
 Db 192 agtcagcctccctggcggtgtgaaagagactcggagtcgctctccaaagtgc 251
 QY 121 cgcgtgagtgagctctcaccctcagcccaatgagcctcttcgggctctcctgctg 180
 Db 252 cgcgtgagtgagctctcaccctcagcccaatgagcctcttcgggctctcctgctg 311
 QY 181 acatctgcccggccgagagacaggggactcaggggaatcccaacctgagtagtaaa 240
 Db 312 acatctgcccggccgagagacaggggactcaggggaatcccaacctgagtagtaaa 371
 QY 241 ttccagttttccagcaacaggaacagacgagtagtacaagatcctcagcagtagagaatt 300
 Db 372 ttccagttttccagcaacaggaacagacgagtagtacaagatcctcagcagtagagaatt 431
 QY 301 attactgtgtctactaatggaagtattccagcccaaggtttcctcactatccaaga 360
 Db 432 attactgtgtctactaatggaagtattccagcccaaggtttcctcactatccaaga 491
 QY 361 aatacggcttggtgagattagtagcagtagagagaaaatgatggatacaactaacg 420
 Db 492 aatacggcttggtgagattagtagcagtagagagaaaatgatggatacaactaacg 551
 QY 421 ttgatgaagatttggtctgaagaccagagatgacatatgcaagtattgttga 480
 Db 552 ttgatgaagatttggtctgaagaccagagatgacatatgcaagtattgttga 611
 QY 481 gaagttgaggaacccagtgatggaactatattaggcgctggtggtggtctggtactgta 540
 Db 612 gaagttgaggaacccagtgatggaactatattaggcgctggtggtggtctggtactgta 671
 QY 541 ccaggaaaacagatttctaaagaaaatcaaataggataaagatttgtatctgtagtaata 600
 Db 672 ccaggaaaacagatttctaaagaaaatcaaataggataaagatttgtatctgtagtaata 731
 QY 601 ttctcttgaaacaggggtctcaccactacaacattgtcagccacaaattccagaa 660
 Db 732 ttctcttgaaacaggggtctcaccactacaacattgtcagccacaaattccagaa 791
 QY 661 gctgtgagctctcagtgctaccccttcagcttgcacctggacctgcttaataatgct 720

Db	792	 gcgtgagtccttcagtgctaccccttcagcttgcacatggacgtctctaaatgct	851
Qy	721	ataactgcctttagtagtaccttgaagaccttatctgatattcttgaaccagagagatgagcag	780
Db	852	ataactgcctttagtagtaccttgaagaccttatctgatattcttgaaccagagagatgagcag	911
Qy	781	ttggacttagaagatctatatagggcacaacttggcaacttcttggcaaggctttgtttt	840
Db	912	ttggacttagaagatctatatagggcacaacttggcaacttcttggcaaggctttgtttt	971
Qy	841	ggaagaaaaatccaagatggtggtatctgaaccttctaaacagagagagttaagattatacagc	900
Db	972	ggaagaaaaatccaagatggtggtatctgaaccttctaaacagagagagttaagattatacagc	1031
Qy	901	tgcacacctcgtaacctcttcagtgctcctaaaggaagaaactaaagaaacccgataccatt	960
Db	1032	tgcacacctcgtaacctcttcagtgctcctaaaggaagaaactaaagaaacccgataccatt	1091
Qy	961	tcttggcagggttgtctctcgtttaaactcgttggtggaacttgcctgttgtctccac	1020
Db	1092	tcttggcagggttgtctctcgtttaaactcgttggtggaacttgcctgttgtctccac	1151
Qy	1021	aattgcaatgaattcgaatgtgtccaaagcaagattactataaaaaatcacacaggtctct	1080
Db	1152	aattgcaatgaattcgaatgtgtccaaagcaagattactataaaaaatcacacaggtctct	1211
Qy	1081	cagttgagaccaaagaccggttcaggggatgtgcacaaactcactcacacgagtgagccctg	1140
Db	1212	cagttgagaccaaagaccggttcaggggatgtgcacaaactcactcacacgagtgagccctg	1271
Qy	1141	gagcacatgaggagtgtgactgtgtgcagagggagcacagaggagtagcccatcac	1200
Db	1272	gagcacatgaggagtgtgactgtgtgcagagggagcacagaggagtagcccatcac	1331
Qy	1201	caccagcagctcttgcacagagctgtgcagtttlygcagtggtgattctattagaagac	1260
Db	1332	caccagcagctcttgcacagagctgtgcagttt---gcagtggtgattctattagaagac	1387
Qy	1261	gtatggttatctccatctcttaactcagtttgcctcgaagacatttcaatctcagg	1320
Db	1388	gtatggttatctccatctcttaactcagtttgcctcgaagacatttcaatctcagg	1447
Qy	1321	atttcagtgcatcttgaagaggagacatcaacacagaattaggagttgtgcaacagctc	1380
Db	1448	atttcagtgcatcttgaagaggagacatcaacacagaattaggagttgtgcaacagctc	1507
Qy	1381	ttttgagaggagcctaaagacagagagaaaaaggttctcaatcgttggaagaaaaattaaa	1440
Db	1508	ttttgagaggagcctaaagacagagagaaaaaggttctcaatcgttggaagaaaaattaaa	1567
Qy	1441	tgttgattataatagatcacacagctagtttcagagttaccatgtacgtattccactagct	1500
Db	1568	tgttgattataatagatcacacagctagtttcagagttaccatgtacgtattccactagct	1627
Qy	1501	gggttctgtatttcagttcttcgtacacggttaggttaaagtcatgacagagaaaaaac	1560
Db	1628	gggttctgtatttcagttcttcgtacacggttaggttaaagtcatgacagagaaaaaac	1687
Qy	1561	tgtgcaagtggacactgtattccgttgccttgccttaactctaaagctccatgtcctgggc	1620
Db	1688	tgtgcaagtggacactgtattccgttgccttgccttaactctaaagctccatgtcctgggc	1747
Qy	1621	ctaaaaatcgataaaatcggatttttttttttttttttttttttgcataatccatatataa	1680
Db	1748	ctaaaaatcgataaaatcggatttttttttttttttttttttttgcataatccatatataa	1806
Qy	1681	accagaacattctatgtactacaaacctgtttttaaagaaagaaactatgttgcctatgaat	1740
Db	1807	accagaacattctatgtactacaaacctgtttttaaagaaagaaactatgttgcctatgaat	1866
Qy	1741	taaaactgtgtctgtctgatagga	1764

Db 1867 taaactgtgtcatgctgtagga 1890

RESULT 15
AAZ48599

ID AAZ48599 standard; CDNA; 2858 BP.

XX AC AAZ48599;

XX 25-APR-2000 (first entry)

XX Bone morphogenic protein (BMP) encoding CDNA (clone HETAB62).

XX Bone morphogenic protein; BMP; cytostatic; osteopathic; angiogenic;
KW vulnary; bone disorder; osteoarthritis; cartilage defect; human;
KW tissue repair; gene therapy; ss.

XX Homo sapiens.

Key	Location/Qualifiers
FT CDS	327..1364
FT FT	/*tag= a
FT FT	/product= "bone morphogenic protein"
FT FT	327..392
FT FT	/*tag= b
FT mat_peptide	393..1361
FT FT	/*tag= c
XX WO200004183-A1.	
XX 27-JAN-2000.	
XX 14-JUL-1999; 99WO-US15783.	
XX 15-JUL-1998; 98US-0092922.	
XX (HUMA-) HUMAN GENOME SCI INC.	
XX Ruben SM, Young PE;	
XX WPI; 2000-182442/16.	
XX P-PSDB; AAY59285.	
XX Novel cDNA encoding human bone morphogenic proteins, vectors, host cells and methods of recombinant production, useful for diagnosis and treatment of, e.g. Bone disorders	
XX Claim 1; Page 181-182; 187pp; English.	
XX The invention provides novel human bone morphogenic proteins (BMP) and nucleic acids encoding the BMPs. The BMP polypeptides can be expressed by standard recombinant methodology. Determining the presence or absence of a mutation in the polynucleotides or determining the presence or amount of expression of the polypeptides is useful for diagnosing a pathological condition or a susceptibility to a pathological condition in a subject. The polynucleotides can also be used to prevent, treat or ameliorate a medical condition. The proteins are useful for diagnosis and/or treatment of diseases associated with BMPs, in particular bone disorders (e.g. osteoarthritis, cartilage defects and tissue repair), and in particular for stimulation of angiogenesis. The polynucleotides are useful as reagents for differential identification of tissues or cell types present in biological samples. The polynucleotides can be used in gene therapy to promote the growth of endothelial cells. The present sequence represents a CDNA (clone HETAB62) encoding a BMP of the invention.	
XX Sequence 2858 BP; 851 A; 557 C; 591 G; 859 T; 0 other;	

Query Match 98.0%; Score 1729.2; DB 21; Length 2858;
Best Local Similarity 99.5%; pred. No. 0;
Matches 1756; Conservative 0; Mismatches 3; Indels 5; Gaps 2

3

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 00:41:33 ; Search time 97.06 Seconds
(without alignments)
3366.982 Million cell updates/sec

Title: US-09-457-066-1

Perfect score: 1764

Sequence: 1 attatgtgaaactaccctg.....cttgtctgctgatagga 1764

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630164 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA:*
- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	42.6	2.4	7218	1	US-08-232-463-14
2	37.2	2.1	3098	1	Sequence 14, Appl
3	37.2	2.1	3098	1	Sequence 1, Appl
4	37.2	2.1	3098	1	Sequence 1, Appl
5	37.2	2.1	3098	1	Sequence 1, Appl
6	37.2	2.1	3098	3	Sequence 1, Appl
C 7	35.6	2.0	3581	2	Sequence 1, Appl
C 8	35.2	2.0	29604	3	Sequence 207, App
C 9	35	2.0	1931	3	Sequence 2, Appl
C 10	34	1.9	1959	4	Sequence 1, Appl
11	33.8	1.9	10684	3	Sequence 3, Appl
12	33.6	1.9	1992	1	Sequence 6, Appl
C 13	33.6	1.9	2798	2	Sequence 6, Appl
14	33.2	1.9	528	3	Sequence 24, Appl
15	33.2	1.9	528	4	Sequence 24, Appl
16	33.2	1.9	528	4	Sequence 24, Appl
17	33.2	1.9	528	4	Sequence 24, Appl
C 18	33.2	1.9	1841	5	Sequence 24, Appl
C 19	33	1.9	1875	5	Sequence 1, Appl
C 20	32.8	1.9	4046	1	Sequence 1, Appl
C 21	32.8	1.9	4046	1	Sequence 1, Appl
C 22	32.8	1.9	56516	2	Sequence 1, Appl
23	32.6	1.8	519	3	Sequence 67, Appl
24	32.6	1.8	519	4	Sequence 67, Appl
25	32.6	1.8	519	4	Sequence 67, Appl
26	32.6	1.8	519	4	Sequence 67, Appl
27	32.4	1.8	7218	1	Sequence 14, Appl

28	32.2	1.8	7244	4	US-08-378-313-26	Sequence 26, Appl
C 29	32	1.8	2007	3	US-08-747-221B-36	Sequence 36, Appl
30	32	1.8	2007	3	US-08-747-221B-38	Sequence 38, Appl
31	31.8	1.8	1519	1	US-07-971-759-19	Sequence 19, Appl
C 32	31.8	1.8	2555	2	US-08-693-457-3	Sequence 3, Appl
C 33	31.8	1.8	2555	4	US-09-265-731-3	Sequence 3, Appl
C 34	31.8	1.8	5852	1	US-07-867-106-2	Sequence 2, Appl
C 35	31.6	1.8	857	3	US-08-460-040-1	Sequence 1, Appl
C 36	31.6	1.8	1172	1	US-07-945-288-9	Sequence 9, Appl
C 37	31.6	1.8	1172	1	US-08-462-831-9	Sequence 9, Appl
C 38	31.6	1.8	1172	1	US-08-461-809-9	Sequence 9, Appl
C 39	31.6	1.8	1172	1	US-08-461-441-9	Sequence 9, Appl
C 40	31.6	1.8	1172	5	PCT-US93-08518-9	Sequence 9, Appl
C 41	31.6	1.8	1524	1	US-08-197-792-34	Sequence 34, Appl
C 42	31.6	1.8	1524	1	US-08-459-850-34	Sequence 34, Appl
C 43	31.6	1.8	1524	1	US-08-459-214-34	Sequence 34, Appl
C 44	31.6	1.8	1797	3	US-08-946-026-13	Sequence 13, Appl
45	31.2	1.8	1411	4	US-09-387-574-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; CLONE: ptzgt-F1s
US-08-232-463-14

```

; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 879..2360
US-08-447-500-1

Query Match          2.1%; Score 37.2; DB 1; Length 3098;
Best Local Similarity 54.3%; Pred. No. 0.59;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps

Qy 1296 gcttcaaggagcccttcattcatcttcaggattttacsgtgcattctgaaagaggagacatcaaac 1355
Db 572 GTTTGAATTTCTTTTTTCTGGAGAATTTTGTGTGCACGAGGAAAGAGACGACGAAGA 631
Qy 1356 agaatlaggagtgtgcaacagctcttttgagagagagcctaaaggacagagagaaaaggt 1415
Db 632 AAAAGTTGAACACGACCCACATATATGGAACGTGTTGAAATACAAAGAGAGAAGGTT 691
Qy 1416 ctccaatcgtggaaagaa 1433
Db 692 CGACACTCGAGGAAGCA 709

RESULT 3
US-08-454-097-1
; Sequence 1, Application US/08454097
; Patent No. 5686412
; GENERAL INFORMATION:

```

TITLE OF INVENTION: PROTEIN KINASES
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 233 South Wacker Drive, 6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/454,097
 FILING DATE: 30-MAY-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/185,359
 FILING DATE: 21-JAN-1994
 APPLICATION NUMBER: US 08/008,001
 FILING DATE: 21-JAN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/728,783
 FILING DATE: 03-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5686412and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/31853
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-474-6300
 TELEFAX: 312-474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3098 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:
CLONE: Protein Kinase
FEATURE:
NAME/KEY: CDS
LOCATION: 879..2360
US-08-454-097-1

Query Match 2.1%; Score 37.2; DB 1; Length 3098;
Best Local Similarity 54.3%; Pred. No. 0.59;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1296 gcttcaaggaccttctcattcaggtattacagtgattctgaaagagagagacatcaaac 1355
Db 572 GTTTGAATTTCTTTTCTGGAGAATTTTGGTGCACGAGGAAAGAGAGAGAGAA 631
QY 1356 agaattaggagttgtgaacagctcttttgagagagagcctaaagagagagagaaagt 1415
Db 632 AAAAGTTGAACACGACCATATATGGAACGTGTTGAAATACAAAGAGAGAAAGGTT 691
QY 1416 cttcaatcgtgaaagaa 1433
Db 692 CGACACTCGAGGAAAGCA 709

RESULT 4
US-08-447-408-1
Sequence 1, Application US/08447408
Patent No. 5703377
GENERAL INFORMATION:
APPLICANT: Hoekstra, Merl F.
TITLE OF INVENTION: TYROSINE KINASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,408
FILING DATE: 03-JUL-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/728,783
FILING DATE: 03-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1318
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3098 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Tyrosine Kinase
FEATURE:
NAME/KEY: CDS
LOCATION: 879..2364
US-08-447-408-1

Query Match 2.1%; Score 37.2; DB 1; Length 3098;
Best Local Similarity 54.3%; Pred. No. 0.59;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1296 gcttcaaggaccttctcattcaggtattacagtgattctgaaagagagagacatcaaac 1355
Db 572 GTTTGAATTTCTTTTCTGGAGAATTTTGGTGCACGAGGAAAGAGAGAGAGAA 631
QY 1356 agaattaggagttgtgaacagctcttttgagagagagcctaaagagagagagaaagt 1415
Db 632 AAAAGTTGAACACGACCATATATGGAACGTGTTGAAATACAAAGAGAGAAAGGTT 691
QY 1416 cttcaatcgtgaaagaa 1433
Db 692 CGACACTCGAGGAAAGCA 709
RESULT 5
US-08-453-866-1
Sequence 1, Application US/08453866
Patent No. 5756289
GENERAL INFORMATION:
APPLICANT: Hoekstra, Merl F.
TITLE OF INVENTION: PROTEIN KINASES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,866
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,001
FILING DATE: 20-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-2458
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3098 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Protein Kinase
FEATURE:
NAME/KEY: CDS
LOCATION: 879..2360
US-08-453-866-1

Query Match 2.1%; Score 37.2; DB 1; Length 3098;
Best Local Similarity 54.3%; Pred. No. 0.59;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1296 gcttcaaggaccttctcattcaggtattacagtgattctgaaagagagagacatcaaac 1355
Db 572 GTTTGAATTTCTTTTCTGGAGAATTTTGGTGCACGAGGAAAGAGAGAGAGAA 631

```

Db      632 AAAAGTTGAAACACGACCACATATATGGACGTGTTGAATCAACAAGAGAAGAGGTT    691
Qy      1416 cttcaatcgtggaagaa 1433
                        I   III   II   I
Db      692 CGACACTCGAGGAAGCA 709
                                I   III   II   I

RESULT          7
US-08-738-349-1/c
; Sequence 1, Application US/08738349
; Patent No. 5869638
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Okazaki, Makoto
; APPLICANT: Kawai, Shinji
; APPLICANT: Tsujimura, Atushi
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Cadherin-Like Protein and
; TITLE OF INVENTION: Process for Its Production
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,349
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/364,439
; FILING DATE:
; APPLICATION NUMBER: US 08/112,061
; FILING DATE: 26-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. P.
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 02481.1323-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN: osteoblastic cell line MC3T3E1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 284..2671
US-08-738-349-1

```

```
Query Match      2.0%; Score 35.6; DB 2; Length 3581;
Best Local Similarity 64.8%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 29; Indels 0; Gaps 0
```


Search completed: June 30, 2001, 01:49:04
Job time: 4051 sec

US-09-265-315-24
 ; Sequence 24, Application 05/09265315
 ; Patent No. 6187541
 ; GENERAL INFORMATION:
 ; APPLICANT: Benton, Bret
 ; APPLICANT: Lee, Ving J.
 ; APPLICANT: Malouin, Francois
 ; APPLICANT: Martin, Patrick K.
 ; APPLICANT: Schmid, Molly B.
 ; APPLICANT: Sun, Dongxu
 ; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
 ; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
 ; NUMBER OF SEQUENCES: 111
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; STREET: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071-2066
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; MEDIUM TYPE: storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0
 ; SOFTWARE: Word Perfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/265,315
 ; FILING DATE: March 9, 1999
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/714,918
 ; FILING DATE: September 13, 1996
 ; APPLICATION NUMBER: 60/009,102
 ; FILING DATE: December 22, 1995
 ; APPLICATION NUMBER: 60/003,798
 ; FILING DATE: September 15, 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wardburg, Richard J.
 ; REGISTRATION NUMBER: 32,327
 ; REFERENCE/DOCKET NUMBER: 240/247
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 528 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-265-315-24

	Query Match	Best Local Similarity	Score 33.2;	DB 4;	Length 528;
	Matches 71;	Conservative 0;	Mismatches 0;	Indels 71;	Gaps 0;
QY	300	tattactgtctactaaatgaagtattcacagcccagaatttcttcatacttatccaag	359		
Ddb	343	TAAAGTGTTCCTATTATTCACCAATTACATTTGAATNGNCTCGTNAGNCATTTGTAAG	402		
QY	360	aatatcggtcttggttagagattagcagtagagaaaatgtaaggatacaacttac	419		
Ddb	403	AGATNCGGGCATAAATTTTGTCGCCACAATCAATTTGGTATTTCTTGCTACTACGGCTTAC	462		
QY	420	gtttgatgaagattgggcctt	441		
Ddb	463	GGTGNATTAAATACCNTGGNTT	484		

3